Numero 304
Dicembre 1993
Anno XXVI
Volume LI

La simulazione dei danni cerebrali

Se si addestra alla lettura una rete artificiale di neuroni e poi la si danneggia, gli errori che essa commette sono sorprendentemente simili a quelli riscontrabili in pazienti affetti da specifiche lesioni cerebrali

di Geoffrey E. Hinton, David C. Plaut e Tim Shallice

Nel 1944 un soldato, G.R., fu colpito alla testa da una pallottola. Sopravvisse, ma con una strana disfunzione: riusciva a leggere e a capire senza difficoltà certe parole, mentre molte altre gli creavano problemi. Leggeva la parola *antique* (oggetto d'antiquariato) come «*vase*» (vaso) e *uncle* (zio) come «*nephew*» (nipote).

La lesione ha avuto effetti disastrosi per il paziente, ma ha fornito informazioni preziosissime ai ricercatori che studiano i meccanismi attraverso i quali il cervello comprende il linguaggio scritto. Quando funziona perfettamente, il sistema che traduce in suoni del parlato i segni scritti sulla carta ci rivela ben poco della sua struttura interna, ma quando esso ha subito qualche lesione, il carattere del tutto peculiare della disfunzione che ne risulta può offrire indizi essenziali per individuare l'architettura originaria del sistema intatto.

Negli ultimi anni è stato possibile perfezionare le tecniche di simulazione al calcolatore delle funzioni cerebrali, al punto da poterle utilizzare per costruire modelli dei percorsi di elaborazione delle informazioni. Abbiamo scoperto che, quando li si danneggia deliberatamente, i sistemi artificiali possono presentare un comportamento che ricorda da vicino i sintomi mostrati da esseri umani che abbiano subito lesioni cerebrali. In effetti, la costruzione di un modello che commette gli stessi errori di un essere umano che abbia subito danni cerebrali ci fa sperare di avere imboccato la strada giusta verso la comprensione del funzionamento del cervello.

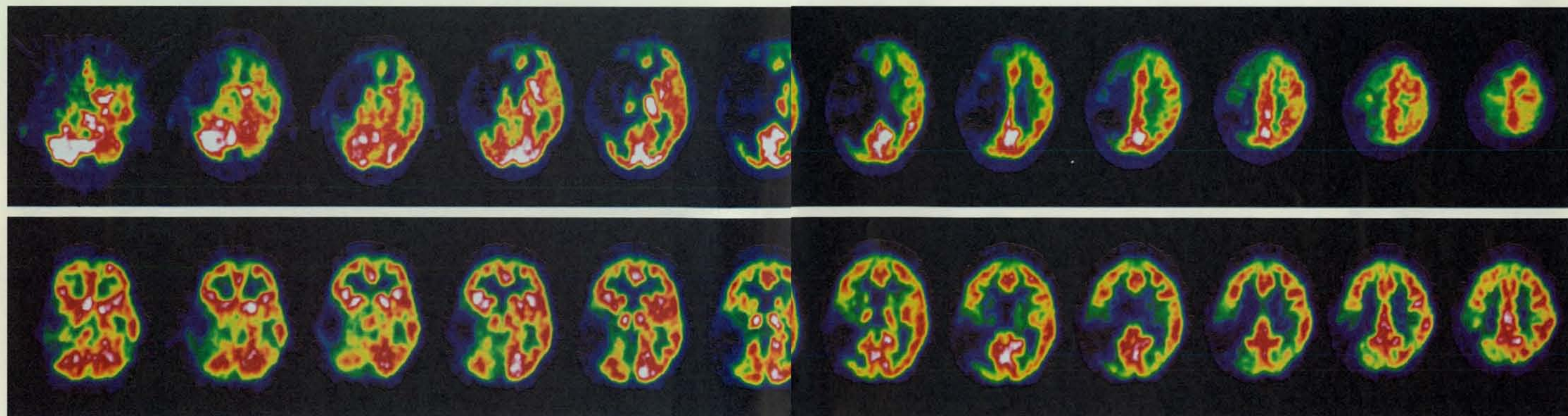
Siamo ben lontani dall'aver realizzato modelli al calcolatore che presentino anche una minima parte delle capacità del cervello umano. Ciononostante, i nostri risultati ci hanno fin qui consentito di ottenere informazioni insperate sui modi in cui il cervello ricava, da una successione di forme di caratteri, il significato di una parola.

Analizzando, nel 1966, le disfunzioni di cui soffriva G.R., John C. Marshall e Freda Newcombe dell'Università di Oxford riscontrarono uno schema molto particolare di deficit di lettura. Oltre a molti errori di semantica, G.R. commetteva anche alcuni errori di visione, leggendo *stock* (provvista) come «*shock*» e *crowd* (folla) come «*crown*» (corona). Spesso la parola che leggeva assomigliava, per forma e significato, a quella corretta: per esempio, vedeva *wise* (saggio) e leggeva «*wisdom*» (saggezza).

Una serie di test molto particolareggiati rivelò che G.R. era in grado di leggere sostantivi riferiti a oggetti concreti, come *table* (tavolo), molto più facilmente di sostantivi astratti come *truth* (verità). Leggeva abbastanza bene i sostantivi (correttamente al 46 per cento), male gli aggettivi (al 16 per cento), ancora peggio i verbi (al 6 per cento); i risultati peggiori erano quelli che riguardavano le particelle funzionali, come le preposizioni (lette correttamente solo nel 2 per cento dei casi). Infine, gli risultava del tutto impossibile leggere successioni di lettere simili a parole, ma prive di senso, come *mave* o *nust*.

In seguito sono stati studiati oltre 50 altri pazienti che commettono errori semantici nella lettura a voce alta, e pressoché tutti hanno presentato la stessa curiosa combinazione di sintomi. Nel 1973 Marshall e la Newcombe hanno descritto due tipi contrastanti di dislessia acquisita. I pazienti affetti dalla cosiddetta dislessia superficiale leggono erroneamente parole dalla pronuncia irregolare, pronunciandole in maniera più aderente alle regole dell'inglese: per esempio, potrebbero leggere *yacht* come se fosse scritto «*yatched*». Un paziente affetto da dislessia «profonda», come G.R., potrebbe invece leggere *yacht* come «*boat*» (imbarcazione).

Per spiegare l'esistenza di questi due tipi di dislessia, Marshall e la Newcombe hanno formulato l'ipotesi che l'informazione elaborata nella lettura normale si propaghi lungo due percorsi distinti e complementari. Coloro che soffrono di dislessia superficiale conservano la via fonologica, che si basa sulla comune corrispondenza fra ortografia e suono. Nel caso della dislessia profonda, invece, si conserva la via semantica, che consente di derivare il significato di una parola direttamente dalla sua forma visiva (quando è possibile derivarlo). Una persona che legga ad alta voce una serie di parole scritte utilizzando solo la via semantica deriva la pronuncia interamente dal significato.

Secondo Marshall e la Newcombe, gli errori commessi dai pazienti sofferenti di dislessia profonda rifletterebbero il funzionamento della via semantica presa a sé. Dati empirici ottenuti successivamente fanno pensare che questa spiegazione sia troppo semplificata, ma l'idea di una via semantica è tuttora generalmente accettata. Oggi sembra probabile che nella dislessia profonda non solo si perda la via fonologica, ma risulti danneggiato anche qualche tratto della via semantica.

Oggi è possibile simulare con reti neurali artificiali le lesioni alle aree di elaborazione del linguaggio che si hanno nei pazienti con dislessia acquisita. (Le immagini sono state ottenute, mediante tomografia a emissione di positroni, da Cathy J. Price e colleghi alla MRC Cyclotron Unit di Londra e misurano l'attività del cervello in sezioni orizzontali successive, a partire dall'alto. Bassi livelli di attività appaiono in blu, livelli elevati in bianco.) Un paziente (*immagini in alto*) ha perso quasi del tutto le funzioni dell'emisfero sinistro della corteccia cerebrale, fatta eccezione per le zone posteriori. L'altro ha subìto lesioni ai lobi parietale e temporale dell'emisfero sinistro, aree ritenute fondamentali per l'elaborazione del linguaggio.

Due diversi percorsi, all'interno del cervello, sono responsabili dell'elaborazione mentale e della pronuncia delle parole scritte. Il primo percorso (la via fonologica) deriva la pronuncia di una parola dall'ortografia, l'altro (la via semantica) la deriva dal significato. I pazienti che sono affetti da dislessia profonda hanno perduto completamente la via fonologica e hanno subito danni anche alla via semantica.

L'ipotesi che la lettura sia correlata a più percorsi di elaborazione, uno dei quali può essere danneggiato mentre gli altri sono intatti, si è rivelata feconda per la classificazione dei pazienti, ma meno utile per capire la natura precisa dei loro problemi. Max Coltheart della Macquarie University in Australia ed Eleanor M. Saffran della Temple University hanno ipotizzato, per esempio, che la lettura dei pazienti affetti da dislessia profonda possa presentare forti somiglianze con quella di persone che hanno solo l'emisfero cerebrale destro funzionante.

La spiegazione, però, non ci è di grande aiuto per capire l'andamento caratteristico degli errori che si presentano normalmente nella dislessia acquisita. Una spiegazione particolareggiata dell'origine di questi errori e del perché formino schemi costanti richiede un modello di come vengano elaborate le informazioni in ciascun percorso, che ci dica anche come si guasta il meccanismo di elaborazione quando i circuiti nervosi sono danneggiati. Gli psicologi usano spesso descrizioni astratte, di tipo algoritmico, dei modi in cui il cervello manipola le informazioni. Ovviamente non è possibile sottoporre questi schemi astratti agli stessi tipi di lesioni che le cellule cerebrali possono subire.

Per questo motivo ci siamo rivolti alle reti neurali, simulazioni elettroniche idealizzate di gruppi di neuroni. Abbiamo messo a punto reti che svolgono il ruolo della via semantica e poi abbiamo eliminato selettivamente alcuni collegamenti fra neuroni per vedere come cambiasse il loro comportamento. Qualche anno fa abbiamo progettato una semplice rete per simulare la via semantica e abbiamo trovato che, danneggiandone una parte qualsiasi, la rete poteva riprodurre numerosi sintomi della dislessia profonda. Da allora abbiamo costruito modelli più particolareggiati per capire quali aspetti delle architetture a rete neurale fossero responsabili di quel comportamento. Abbiamo anche ampliato la nostra ricerca, in modo da spiegare ulteriori sintomi della dislessia profonda.

I nostri modelli della via semantica sono costituiti da unità interconnesse che rappresentano i neuroni. Ciascuna di queste unità ha un livello di attività (compreso fra 0 e 1) dipendente dai segnali che riceve in ingresso da altri neuroni. I collegamenti fra le unità hanno un peso modificabile, che specifica in quale misura l'uscita di una di esse si rifletterà sull'attività dell'unità a cui invia il suo segnale. Unitamente alla configurazione dei collegamenti fra i vari neuroni, questi pesi determinano quali elaborazioni la rete esegua.

Alla base delle reti neurali artificiali vi è un neurone idealizzato, un dispositivo che somma gli ingressi pesati, ricevuti da altri neuroni (*in basso*), e genera un livello di attivazione compreso fra 0 e 1; quindi passa questa attivazione (lungo connessioni pesate) ad altri neuroni. Il comportamento della rete è determinato dall'insieme dei pesi e delle connessioni.

La prima versione della nostra rete era costituita da tre gruppi di unità: quelle «grafemiche», ciascuna delle quali rappresentava una particolare lettera in una posizione specifica all'interno di una parola; quelle «sememiche» che rappresentavano i significati delle parole; e uno strato di unità intermedie che rendevano possibile l'apprendimento di associazioni complesse. Una rete generale richiederebbe 26 unità grafemiche per ciascuna posizione all'interno di una parola: per poter ridurre il numero delle unità necessarie abbiamo utilizzato un vocabolario semplificato. Le unità grafemiche in prima posizione erano tutte consonanti, per esempio, e quelle in seconda posizione tutte vocali.

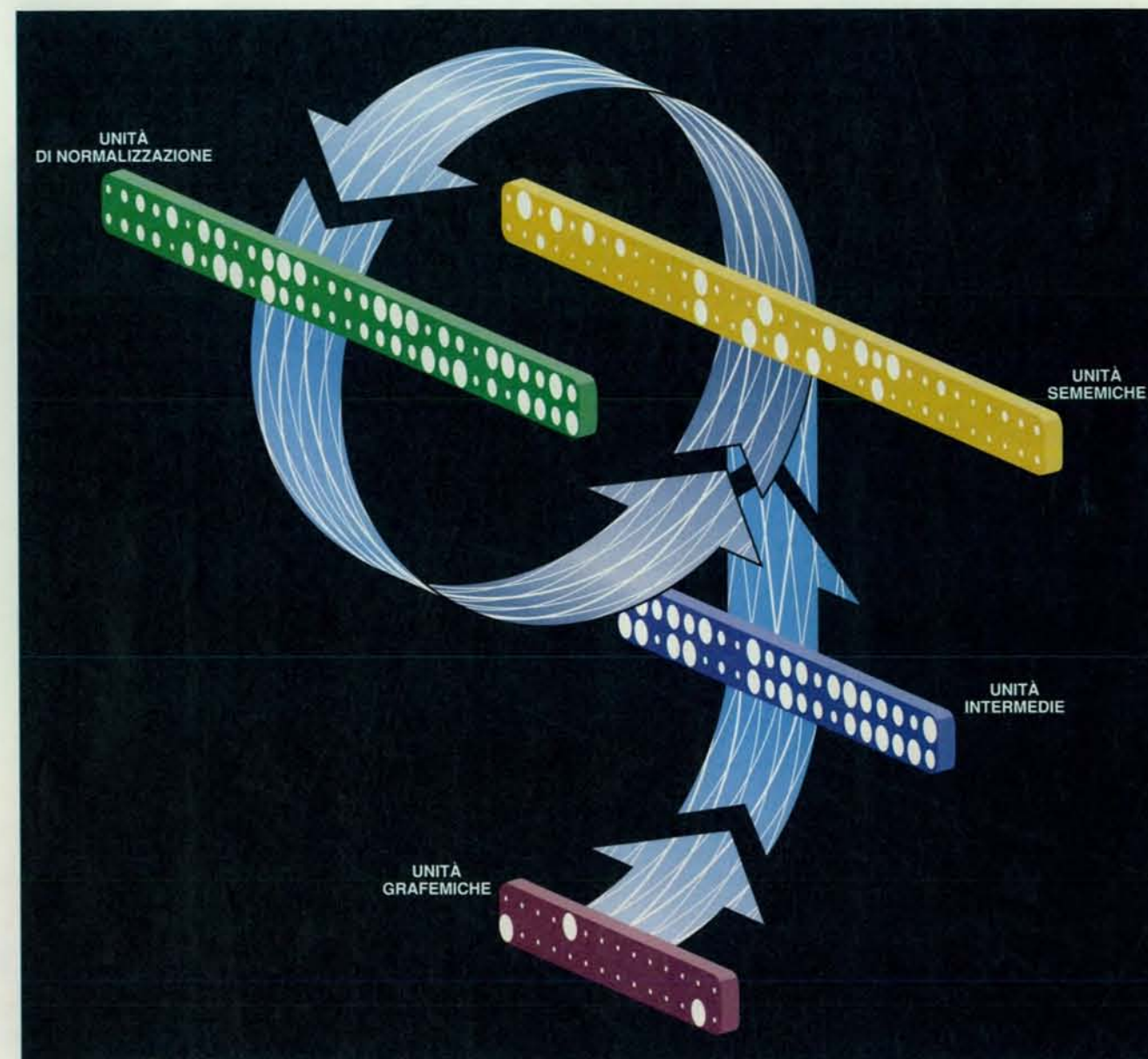
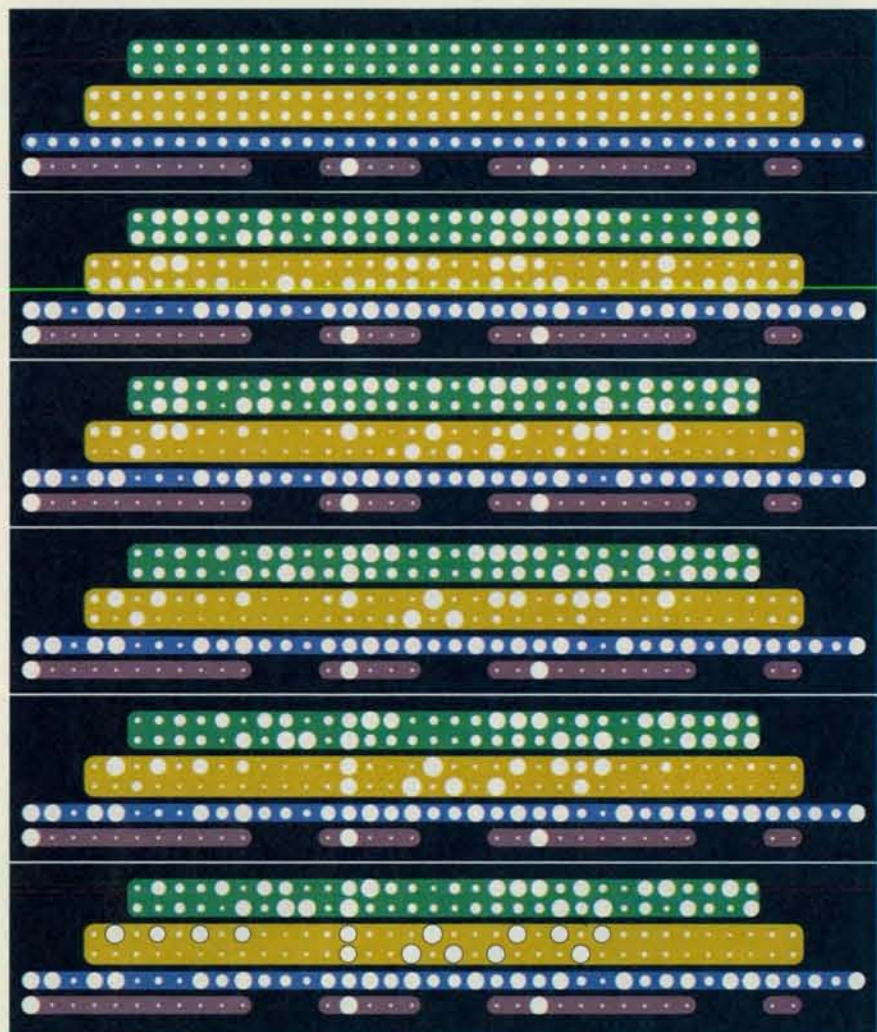
Le unità sememiche non corrispondono direttamente ai significati delle singole parole, bensì a tratti semantici che descrivono l'entità in questione. La parola *cat* (gatto) attiva unità come «mammifero», «dotato di zampe», «morbido» e «orgoglioso». Le unità che rappresentano tratti semantici come «trasparente», «di gusto forte», «parte di arto» o «fatto di legno» restano inattive. La nostra rete ha 68 unità sememiche che rappresentano attributi fisici e funzionali nella definizione di una parola. Ogni parola che abbiamo scelto era rappresentata da una diversa combinazione di unità sememiche attive e inattive.

Perché la nostra rete neurale producesse la giusta configurazione di tratti semantici per ciascuna parola, abbiamo dovuto attribuire il valore opportuno al peso di ciascuna connessione. I pesi però non vengono assegnati manualmente, bensì attraverso un processo di apprendimento: un algoritmo per la programmazione delle reti neurali. Per insegnare un certo compito a una rete, si inizia con un insieme casuale di pesi e si presenta varie volte alla rete un «insieme di addestramento» di configurazioni di ingresso (nel nostro caso, lettere in posizioni ben definite). Dopo ogni sessione di addestramento l'algoritmo modifica i pesi in modo da minimizzare la differenza fra l'uscita della rete e la risposta «corretta».

Già negli anni cinquanta si sapeva come regolare i pesi in semplici reti neurali a due strati, ma l'addestramento di reti con un numero più elevato di strati presenta maggiori difficoltà. In particolare, non è immediatamente ovvio come impostare i pesi delle connessioni tra unità di ingresso e unità intermedie perché non vi è modo di determinare, a priori, quali unità intermedie debbano essere attive per determinati ingressi e uscite.

Nel corso degli anni ottanta, tuttavia, sono stati sviluppati parecchi metodi differenti per l'addestramento delle reti a più strati. Questi metodi assegnano le variazioni nei pesi delle connessioni di ciascuno strato in funzione del loro contributo all'errore dell'uscita. Al crescere del numero dei cicli di addestramento, i pesi risultanti convergono a un insieme di valori che definisce una rete capace di produrre risultati corretti. In base ai pesi assegnati inizialmente a caso, l'apprendimento può avere come esito molti insiemi diversi di pesi, ciascuno dei quali fa sì che la rete produca risposte corrette ai segnali forniti in ingresso durante l'addestramento. (Per ulteriori particolari sulla procedura di apprendimento, si veda l'articolo *L'apprendimento delle reti artificiali di neuroni* di Geoffrey E. Hinton in «Le Scienze» n. 291, novembre 1992.)

In teoria, queste procedure di apprendimento possono impantanarsi in qualche minimo locale, cioè in una configurazione di pesi non corretta, ma tale per cui qualsiasi piccola variazione non farebbe altro che peggiorare gli errori commessi dalla rete. In pratica, però, una rete apprende per lo più soluzioni quasi ottime. Inoltre alcune procedure di apprendimento sono più plausibili, dal punto di vista biologico, di altre, ma non sembra che i nostri risultati dipendano dal particolare metodo utilizzato. Pensiamo che, anche se il cervello segue una procedura di apprendimento molto diversa, i circuiti nervosi cerebrali continuino ad assomigliare alla struttura sviluppata dalla nostra rete. Pertanto la nostra spiegazione di quel che succede quando la rete è danneggiata può essere corretta anche se non sono corrette le sue procedure di apprendimento.

La nostra prima rete, dotata di un solo strato intermedio, era in grado di apprendere come far corrispondere tratti semantici a forme di parole, ma non era del tutto soddisfacente. Presentava infatti una forte tendenza a far corrispondere ingressi molto simili (per esempio *cat* e *cot*) con uscite simili, se non sottoposta a un addestramento estremamente lungo. Abbiamo affrontato il problema ag-

Una rete neurale per la lettura ha quattro strati. Il primo reagisce alle lettere di una parola; le connessioni fra le unità di ingresso e le unità intermedie e fra queste e le unità sememiche traducono le forme di parola in rappresentazioni in termini di tratti semantici, come «è grande», «è vivo», «è commestibile». Le unità di normalizzazione sono connesse alle unità sememiche in un anello di retroazione che regola l'uscita di queste ultime in modo da farla corrispondere ai significati delle parole.

I livelli di attivazione dei neuroni della rete variano nel tempo, via via che la rete elabora la parola *bed* (letto). Inizialmente sono attivate, in misura diversa, numerose unità sememiche, ma l'interazione con le unità di normalizzazione rafforza l'attivazione di alcune e indebolisce quella di altre, fino a che l'uscita non converge.

giungendo un altro strato di neuroni «di normalizzazione». Se l'insieme iniziale di connessioni genera ùna risposta inadeguata, le nuove unità la modificano, producendo esattamente i tratti semantici corretti. Il numero dei significati è limitato, sicché il percorso che parte dall'ingresso deve soltanto portare le attività delle unità sememiche più vicine al significato corretto che a tutti gli altri. Le stesse tecniche di apprendimento che funzionano sulle reti con un solo strato intermedio possono regolare l'apprendimento anche di reti che contengono più strati intermedi, e addirittura di reti le cui unità sono collegate a formare circuiti chiusi.

Il modo più naturale di realizzare questo meccanismo di normalizzazione consiste nell'utilizzare un anello di retroazione. L'uscita delle unità sememiche va alle unità di normalizzazione, e l'uscita di queste ultime va agli ingressi delle unità sememiche. Ogni volta che l'attività percorre il ciclo, l'influenza delle unità di normalizzazione sulle unità sememiche (e viceversa) produce una configurazione di tratti semantici più vicina a quella corretta.

L'anello di retroazione introduce una nuova caratteristica nel comportamento della nostra rete neurale. La rete iniziale era statica: un dato valore in ingresso faceva sì che la rete producesse un corrispondente schema in uscita, e questo schema non cambiava fino a che l'ingresso rimaneva costante. L'uscita della nuova rete, tuttavia, è dinamica: si assesta gradualmente su una configurazione stabile.

Di conseguenza abbiamo trovato utile pensare all'uscita della rete non semplicemente come a un elenco di tratti semantici attivi, bensì come a un moto in uno «spazio semantico» pluridimensionale, le cui coordinate sono definite da tutti i tratti semantici che la rete può rappresentare. Ciascun punto dello spazio corrisponde a uno specifico schema di attività fra le unità sememiche, ma solo alcuni di questi schemi sono associati a significati validi. I significati corretti delle parole sono punti nello spazio semantico.

I primi tre strati della rete, visti in questa prospettiva, prendono una forma di parola e la trasformano in una posizione nello spazio semantico. L'attività nello strato di normalizzazione porta poi l'uscita della rete al punto che corrisponde al significato più vicino. La regione circostante ogni parola è quello che in fisica e in matematica si chiama attrattore: ogniqualvolta l'uscita iniziale della rete finisce all'interno di una certa regione, lo stato della rete verrà inesorabilmente trascinato verso una particolare posizione all'interno di quella regione.

Questa idea di uno spazio semantico punteggiato di attrattori che rappresentano i significati di parole si è dimostrata valida per capire come funzioni la nostra rete neurale e come possa compiere gli stessi errori semantici dei dislessici. Se danneggiamo la rete modificando a caso i pesi nel meccanismo di normalizzazione, per esempio, i confini dell'attrattore per ciascuna parola cambieranno. Di conseguenza, se la rete si trova in una regione dello spazio semantico in cui è stata attratta in precedenza verso una certa parola, può darsi che ora venga attratta verso un'altra parola che ha qualche rapporto semantico con la precedente. In alternativa, se danneggiamo il percorso proveniente dall'ingresso, l'uscita iniziale della rete può essere più vicina al significato di una parola semanticamente correlata che non al significato della parola effettivamente presentata.

Questo risultato chiarisce uno dei primi enigmi presentati dalla dislessia profonda: perché lesioni provocate a una qualsiasi parte della via semantica del cervello producano uno schema essenzialmente simile di errori di lettura. I neurologi, e non solo loro, si sono chiesti come sia possibile che una lesione in prossimità dell'ingresso (la parte visiva del sistema di lettura) produca errori semantici. Secondo i nostri modelli, questi errori nascono naturalmente quando i neuroni di normalizzazione usano informazioni semantiche per cercare di dare un senso all'uscita degli stadi danneggiati che li precedono.

Il concetto di attrattore ci aiuta a spiegare anche un'altra anomalia dei dati. Quasi tutti i pazienti che commettono errori semantici compiono anche alcuni errori visivi: confondono una parola come *cat* con una dall'aspetto simile come *cot*. Non commettono però gli errori di pronuncia di chi è affetto da dislessia superficiale («*loave*» per *love* o «*deef*» per *deaf*). Questo collegamento invariabile fra errori semantici ed errori visivi è strano. Alcuni pazienti debbono avere subìto danni solo agli strati più avanzati del sistema di elaborazione, e ci si aspetterebbe intuitivamente che commettano soltanto errori semantici.

Dopo aver realizzato il nostro modello a rete neurale, abbiamo scoperto con grande sorpresa che danni al circuito di normalizzazione semantica talvolta causavano errori visivi. Col senno di poi, possiamo capire perché: gli strati iniziali di una rete non danneggiata possono permettersi di produrre uscite semantiche relativamente simili per le parole *cat* e *cot* perché poi il circuito di normalizzazione orienterà ciascuna verso il suo significato appropriato. Quando però il circuito di normalizzazione è danneggiato e la forma degli attrattori cambia, l'uscita delle unità sememiche può cadere nell'area di influenza dell'attrattore per una parola visivamente simile, ma priva di alcun collegamento semantico.

Questa spiegazione non ci è venuta in mente subito, perché si basa sull'idea che il confine dell'attrattore per *cat* può arrivare molto vicino a quello per *cot* anche se le due parole non hanno parentele semantiche. Ci si aspetterebbe che fra gli attrattori per *cat* e *cot* se ne trovino molti altri per significati diversi. In uno spazio bidimensionale l'intuizione è corretta: se scegliamo a caso 40 punti in modo che rappresentino significati di parole e costruiamo attorno a ciascun punto attrattori piuttosto compatti, gli attrattori per significati diversi non sono mai vicini.

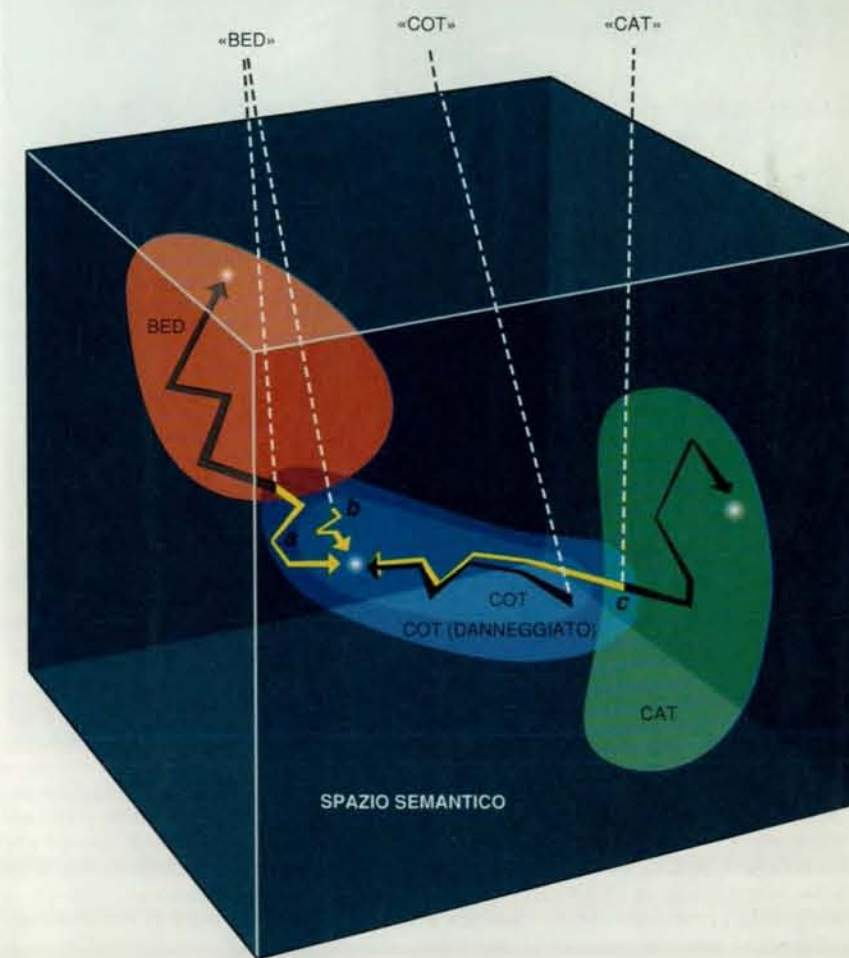
È molto pericoloso, però, dare per scontato che la stessa cosa valga in spazi a molte dimensioni. La nostra rete rappresenta nelle sue unità sememiche 68 tratti semantici, sicché gli attrattori per ciascuna delle sue 40 parole si trovano in uno spazio a 68 dimensioni. Il fatto è che, in 68 dimensioni, il punto medio tra due punti scelti a caso è quasi sicuramente più vicino a ciascuno di quei due punti che non a tutti gli altri 38 punti. Di conseguenza gli attrattori per *cat* e *cot* possono avere una frontiera in comune e non essere separati da altri attrattori. In uno spazio a 68 dimensioni è facile evitare gli ostacoli.

La nostra rete è in grado di simulare le corrispondenze, sia corrette sia derivanti da disfunzioni, fra forme di parole e significati, ma questo non significa che la sua architettura sia l'unica possibile per la via di elaborazione semantica nel cervello. Per determinare l'insieme delle possibili alternative, abbiamo studiato gli effetti di danni su un certo numero di architetture diverse, ciascuna progettata in maniera tale da mettere alla prova un particolare aspetto del progetto della rete originale.

Abbiamo programmato versioni della rete neurale che contenevano connessioni fra le unità sememiche e versioni in cui tali connessioni erano assenti; abbiamo anche programmato alcune reti nelle quali ciascun neurone di uno strato era connesso a tutti i neuroni dello strato successivo e altre in cui ciascun neurone era collegato solo a pochi neuroni dello strato successivo. Abbiamo spostato le unità di normalizzazione in modo che svolgessero il loro lavoro prima delle unità sememiche, e abbiamo combinato le unità di normalizzazione con lo strato intermedio. Abbiamo addirittura variato la disposizione dei neuroni nello strato di ingresso, in modo da modificare la rappresentazione delle parole, e abbiamo aggiunto una rete di uscita che convertiva i significati in successioni di fonemi, affinché il sistema effettivamente parlasse.

Per lo più, i particolari architettonici si sono dimostrati irrilevanti. Il modo specifico di rappresentazione dell'ingresso visivo non è importante, purché le parole che si assomigliano visivamente producano schemi simili di attività nello strato di ingresso. L'unico ingrediente cruciale è l'esistenza di attrattori: se non ci sono unità di normalizzazione a valle del danno, la rete non presenta la configurazione di errori caratteristica della dislessia profonda.

È abbastanza interessante che la nostra rete non solo riproduca gli errori visivi e semantici ovvi della dislessia profonda, ma simuli anche alcune delle caratteristiche più fini di questa disfunzione. Per esempio, qualche volta i pazienti compiono errori «semantici visivamente

Lo spazio semantico ha molte dimensioni, corrispondenti ai tratti semantici (il disegno ne offre una semplice approssimazione tridimensionale). I significati di singole parole sono identificati mediante punti nello spazio semantico. Quando la rete neurale costruita dagli autori legge una parola, l'interazione fra le unità sememiche e quelle di normalizzazione fa sì che una forma di parola che sia stata proiettata in una regione dello spazio semantico, detta attrattore, vicina al significato di una parola (*regioni in colore*) converga a quel particolare significato (*punti*). Se la rete è danneggiata e i confini degli attrattori si sono spostati, è possibile che una parola venga letta erroneamente come se fosse una parola semanticamente simile: per esempio «*cot*» per *bed* (*a*). Si possono verificare errori semantici anche quando il danno fa in modo che una forma di parola sia proiettata su un punto leggermente diverso dello spazio semantico (*b*). Una rete neurale può commettere errori perché all'inizio parole visivamente simili vengono proiettate su punti prossimi dello spazio semantico, anche se i punti di stabilità degli attrattori sui quali le parole vanno a cadere sono distanti (*c*).

determinati», nei quali sembra che una confusione semantica segua una confusione visiva. G.R. leggeva *sympathy* come «*orchestra*» (presumibilmente perché vedeva *symphony*). Anche le nostre reti producono errori di questo tipo: talvolta leggono *cat* come «*bed*» (letto), perché vedono *cot*.

Quando è gravemente danneggiata, la nostra rete presenta anche uno strano effetto. Lo stesso fenomeno si nota in pazienti con una lesione di dimensioni tali che le loro rappresentazioni semantiche sono distorte oltre ogni possibilità di riconoscimento; questi pazienti, di conseguenza, non riescono a trovare una parola. Essi non riescono a identificare la parola che cercano di leggere, ma spesso possono ugualmente stabilire in quale categoria quella parola ricada (per esempio, possono dire se si riferisca a un «animale» o a un «alimento»). In circostanze simili, la nostra rete non si stabilizza più sull'attrattore che corrisponde a una particolare parola: in effetti, è possibile che gli attrattori per molte parole si siano fusi in uno solo. Ciononostante, l'uscita della rete si stabilizza all'interno di un volume più grande dello spazio semantico, entro il quale si trovavano in precedenza la parola corretta e quelle a essa correlate. Di conseguenza, è ancora possibile determinare la categoria della parola.

Un sintomo della dislessia profonda che i nostri modelli all'inizio non riproducevano è la maggiore difficoltà nella lettura delle parole astratte rispetto a quelle concrete. Questo fenomeno sembra essere parte integrante della sindrome perché l'astrattezza (una proprietà semantica) aumenta la probabilità di errori visivi. Inoltre, quando i pazienti commettono questi errori di lettura, il risultato del loro errore in genere è più concreto della parola effettivamente presentata.

Per simulare questo effetto, abbiamo seguito un'idea formulata da Gregory V. Jones dell'Università di Warwick in Inghilterra e da altri, secondo la quale per i pazienti affetti da dislessia profonda le parole concrete sarebbero più facili perché evocano un significato più coerente e dettagliato. In termini della nostra rete, una parola concreta ha più tratti semantici di una parola astratta. Per esempio, la parola *post* (paletto) ha 16 tratti, che vanno da «lunghezza compresa tra un piede e due iarde» a «utilizzato per giochi o ricreazione». Invece, *past* (passato) ha solo due tratti: «ha durata» e «si riferisce a un momento precedente nel tempo». Abbiamo progettato un nuovo vocabolario che contiene 20 coppie di parole di quattro lettere, una concreta e l'altra astratta, dove i due elementi di ciascuna coppia differiscono di una sola lettera. In media, le parole concrete hanno un numero di tratti semantici circa quadruplo di quelle astratte.

Dopo aver addestrato la rete a pronunciare queste parole, abbiamo trovato che lesioni in qualsiasi parte della rete a monte delle unità di normalizzazione riproducevano gli effetti della lettura di parole astratte. Le parole concrete provocano un minor numero di errori perché nei loro schemi di attività semantica vi è maggiore ridondanza. Esiste quindi una struttura più ampia utilizzabile dalle unità di normalizzazione per far sì che la rete converga verso il significato corretto. Le parole astratte, che hanno minore ridondanza nei loro schemi semantici, devono affidarsi maggiormente al percorso «anticipativo», dove le influenze visive sono più forti.

Poiché il corretto riconoscimento delle parole concrete si basa maggiormente sul circuito di normalizzazione, se questo viene gravemente danneggiato si ha un sorprendente rovesciamento: la rete danneggiata legge meno bene le parole concrete, producendo un numero di errori maggiore che con le parole astratte.

Una rete che parla presenta un ulteriore insieme di tre strati oltre a quelli della rete neurale per la lettura. Essa converte le successioni di forme delle lettere in rappresentazioni semantiche e poi proietta queste ultime su successioni di fonemi che possono essere inviate a un sintetizzatore vocale. Questa rete è particolarmente utile perché non impone ai ricercatori di formulare giudizi (potenzialmente viziati) su quale parola eventualmente corrisponda a uno schema perturbato di tratti semantici che viene generato quando la rete è stata danneggiata per simulare la dislessia.

Questo tipo di lesione e questo schema di comportamento sono coerenti con quello che sappiamo dell'unico, enigmatico paziente affetto da «dislessia per le parole concrete», studiato da Elizabeth K. Warrington del National Hospital di Londra. Questo paziente non soltanto incontrava maggiori difficoltà nella lettura delle parole concrete rispetto a quelle astratte, ma riusciva meglio a mettere in corrispondenza parole astratte e immagini. Questa coerenza di comportamento fa pensare che il suo problema si trovasse a livello del sistema semantico.

La nostra spiegazione del tipo di errori caratteristici della dislessia profonda si basa sulle proprietà di una rete neurale che trasforma una rappresentazione (la forma visiva di una parola) in un'altra rappresentazione (un insieme di tratti semantici) a essa correlata in modo arbitrario. Ci si aspetterebbe che simili tipi di errori derivino da danni ad altri processi cognitivi che comportano una trasformazione arbitraria in o da uno spazio semantico. I neuropsicologi, inoltre, hanno già descritto tipi di errori abbastanza simili nella disgrafia profonda, un disordine della scrittura, e nella disfasia profonda, un disordine della ripetizione delle parole.

Questi ulteriori dati ci fanno pensare che il nostro modello possa avere una validità più ampia di quanto inizialmente pensassimo; cosa ancor più importante, segna un successo nell'uso di una nuova tecnica per la comprensione del funzionamento del cervello. Il nostro lavoro differisce da altre spiegazioni della dislessia profonda (e, con qualche eccezione, da altre spiegazioni di fenomeni neuropsicologici in generale) per il tipo di ipotesi che mettiamo in campo. Invece di caratterizzare verbalmente ciascuna componente di un complesso meccanismo neurale e basarci sull'intuizione per dire come un danno ne influenzi il comportamento, simuliamo quel meccanismo, lo danneggiamo e stiamo a vedere che cosa succede. Abbiamo trovato che molte delle nostre aspettative erano sbagliate. Questa scoperta ci fa pensare che simulazioni al calcolatore dettagliate svolgeranno un ruolo determinante nell'approfondire la nostra conoscenza dei modi in cui il cervello elabora normalmente informazioni sul linguaggio e di come questa funzione venga danneggiata da lesioni o malattie.

BIBLIOGRAFIA

HINTON G. E. e SHALLICE T., *Lesioning an Attractor Network: Investigations of Acquired Dyslexia* in «Psychological Review», 98, n. 1, gennaio 1991.

PLAUT D. C. e SHALLICE T., *Deep Dyslexia: A Case Study of Connectionist Neuropsychology* in «Cognitive Neuropsychology», 10, n. 5, ottobre 1993.

Le grandi province magmatiche

Queste immani effusioni laviche, dovute a impulsi di attività magmatica intensi ma geologicamente brevi, potrebbero aver causato modificazioni importanti nelle condizioni climatiche e ambientali dell'intero pianeta

di Millard F. Coffin e Olav Eldholm

Nel campo della ricerca geofisica persino giungere a destinazione qualche volta non è uno scherzo. Mentre si stava dirigendo verso ovest durante una tempesta nell'Oceano Indiano meridionale, la nave da ricerca australiana *Rig Seismic* fu investita così duramente dalle onde che la sua àncora si staccò e sfasciò il castello di prua, aprendo una pericolosa falla. Il proseguimento del viaggio verso il plateau delle Kerguelen, una zona sopraelevata del fondo marino poco a nord dell'Antartide, dovette essere ritardato, mentre l'equipaggio assicurava l'àncora e riparava lo scafo della nave. Solo molti giorni più tardi la spedizione, che comprendeva uno di noi (Coffin), arrivò alle Kerguelen e, dopo aver sistemato gli strumenti per l'analisi della struttura del fondo marino, iniziò la raccolta dei dati. Lo scopo del viaggio era quello di comprendere l'origine e l'evoluzione dell'enorme plateau sottomarino.

In quello stesso anno, il 1985, ma nell'altro emisfero, la nave laboratorio *JOIDES Resolution* andava solcando le tempestose acque dell'Atlantico settentrionale verso il plateau di Vøring, al largo delle coste della Norvegia. Quella spedizione, condotta da uno di noi (Eldholm) e da Jörn Thiede del Centro di ricerca per le scienze oceanografiche di Kiel (Germania), aveva il compito di studiare le strutture geologiche che si formano quando i continenti si separano, dando luogo a un nuovo bacino oceanico. Contrariamente alle aspettative della maggior parte dei ricercatori, il gruppo riuscì a ottenere carote di rocce ignee alla profondità di quasi un chilometro, sotto una coltre di 300 metri di sedimenti marini.

Riconoscemmo così alcune fondamentali analogie tra questi due plateau, nonostante il loro diverso ambiente geologico e le migliaia di chilometri di distanza che li separano. I nostri studi, unitamente ai dati sismici e di perforazione ottenuti da molti altri ricercatori, hanno contribuito a dimostrare che il plateau delle Kerguelen e il margine continentale vulcanico della Norvegia appartengono a una classe di enormi lineamenti magmatici genericamente conosciuti con il nome di grandi province magmatiche. Queste strutture coprono aree che si estendono anche per milioni di chilometri quadrati, tuttavia sembrano essersi formate piuttosto rapidamente in termini geologici.

Gli impulsi di attività eruttiva associati a queste rapide effusioni di lava potrebbero aver avuto effetti profondi sul chimismo e sulla circolazione degli oceani e dell'atmosfera. Alcune delle modificazioni ambientali risultanti potrebbero aver contribuito alle estinzioni in massa, compresa quella dei dinosauri; altri cambiamenti potrebbero viceversa aver dato impulso alla diversità biologica e all'origine di nuove specie.

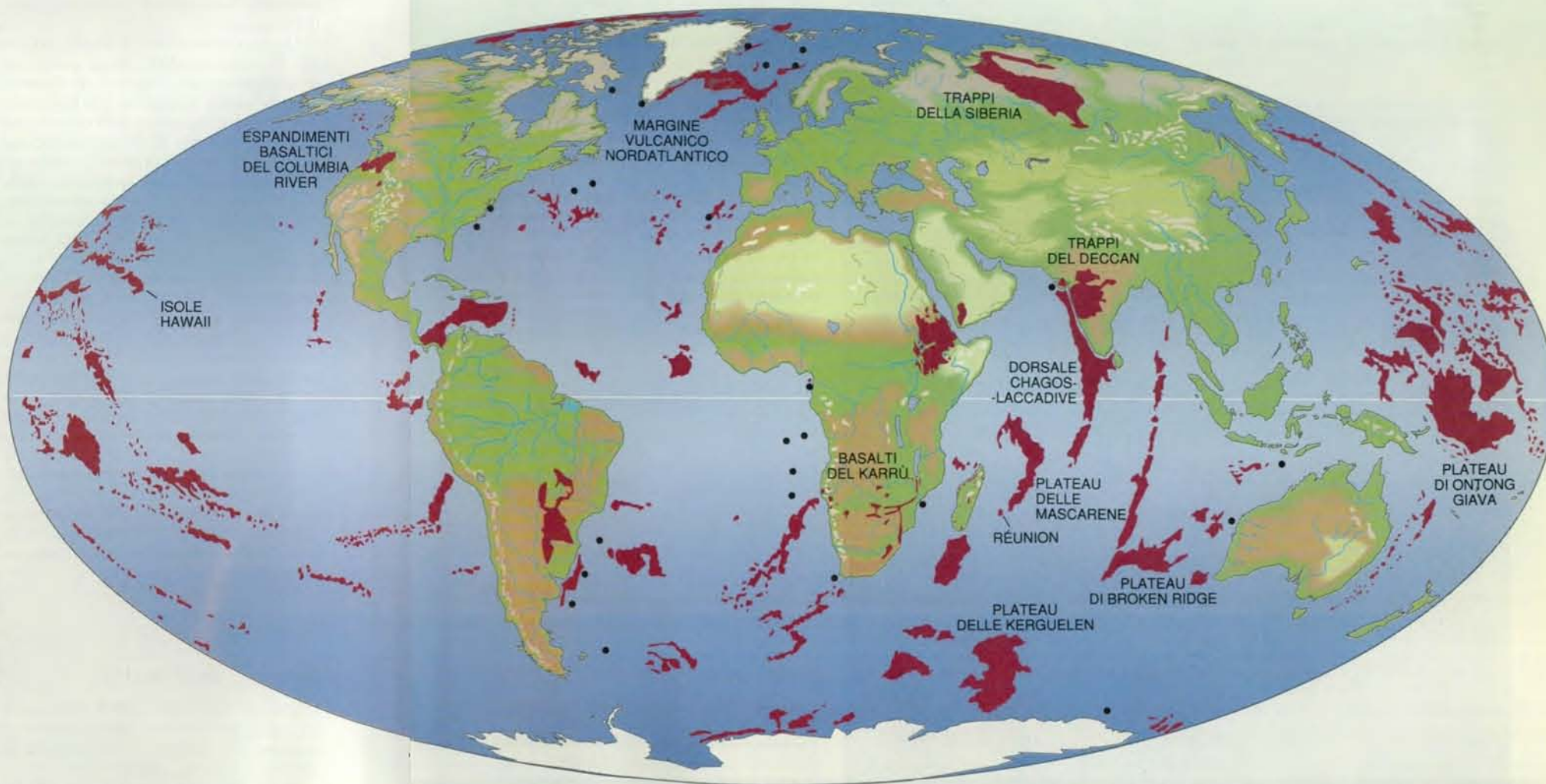
Il riconoscimento delle grandi province magmatiche ha costretto i geofisici a riconsiderare le conoscenze che si ritenevano acquisite sulla struttura interna della Terra. La teoria della tettonica delle zolle descrive in modo soddisfacente la lenta e costante attività vulcanica che avviene lungo le dorsali medio-oceaniche (dove si forma la nuova crosta oceanica) e presso le zone di subduzione (dove le parti più dense e antiche di fondo oceanico sprofondano nell'interno della Terra), ma non può spiegare facilmente le improvvise effusioni necessarie per creare le grandi province magmatiche. Sebbene la composizione delle lave che fuoriescono dalle dorsali medio-oceaniche assomigli in linea di massima a quella delle rocce costituenti le province magmatiche, queste ultime si distinguono per il loro contenuto di elementi in traccia e per la composizione isotopica. Talvolta, inoltre, questi magmi erompono nelle parti interne, normalmente tranquille, delle zolle litosferiche.

Negli anni sessanta il compianto J. Tuzo Wilson dell'Università di Toronto e poi W. Jason Morgan della Princeton University elaborarono una ipotesi che aiutò a comprendere il fenomeno. Essi proposero che il mantello, la vasta zona situata tra la crosta e il nucleo, avesse due modalità diverse di circolazione. La principale consiste in un moto di convezione a grande scala che trascina le zolle litosferiche causando la deriva continentale. Circa un decimo del calore che attualmente fuoriesce dal mantello è prodotto però da pennacchi sottili, che hanno origine in profondità, di materiale più caldo del normale risalente attraverso il mantello stesso. Una volta raggiunta la base della litosfera, questi pennacchi subiscono una decompressione e una fusione parziale, dando luogo a una risalita di magma e a una zona di attività vulcanica duratura meglio nota con il nome di punto caldo. Una parte di questo magma può essere eruttata, e ciò produce un immenso flusso di lava. Diversamente dal meccanismo lento e continuo della deriva continentale e dell'espansione dei fondi oceanici, la risalita di pennacchi del mantello avviene con modalità del tutto episodiche e irregolari.

Tutti i modelli della struttura interna della Terra sono basati su deduzioni; perfino i pozzi più profondi, infatti, superano di poco i 10 chilometri di profondità. Lo studio diretto e indiretto delle grandi province magmatiche assume quindi un'importanza fondamentale per comprendere meglio la natura dei pennacchi del mantello e come essi possano influire sulle condizioni alla superficie terrestre.

La prima, fondamentale osservazione che riguarda le province magmatiche è che esse sono composte da basalto, una roccia comune, ricca di ferro e magnesio. I grandi affioramenti ignei presenti all'interno dei continenti, dove è possibile campionarli facilmente e determinarne la composizione, sono denominati espandimenti basaltici continentali; le province magmatiche presenti lungo i bordi dei continenti sono chiamate margini vulcanici passivi, mentre quelle che si trovano negli oceani sono dette plateau oceanici.

I geologi riconobbero l'esistenza degli espandimenti basaltici continentali alla fine del XIX secolo, quando compresero che molte strutture vulcaniche apparentemente separate fanno parte in realtà di flussi collegati di lave basaltiche. Tra le più spettacolari di queste strutture vi sono i cosiddetti trappi, o tavolati, del Deccan (Deccan Traps) dell'India centro-occidentale. (*Trap* è un termine olandese che significa «scalinata», in riferimento all'aspetto che assumono le lave stratificate ed erose.) Strutture simili, per esempio, sono i basalti del Columbia River nella parte nordoccidentale degli Stati Uniti, le effusioni vulcaniche nordatlantiche delle Isole Britanniche, delle Isole Faer Øer e della Groenlandia, i basalti del Karrù in Africa meridionale. Gli strati di lava che costituiscono un espandimento basaltico producono impilamenti spessi anche molti chilometri; le singole colate possono contenere migliaia di chilometri cubi di roccia ed estendersi per centinaia di chilometri.

Data l'inaccessibilità delle province magmatiche sottomarine, gli studiosi si sono resi conto solo recentemente delle grandi analogie tra queste ultime e gli espandimenti basaltici continentali. Nel 1981 Karl Hinz dell'Istituto federale per

Le grandi province magmatiche (*in rosso*) si trovano nelle zone più disparate del globo. I punti in nero indicano siti lungo i margini continentali dove i profili sismici sembrano rivelare tracce di antichi episodi di vulcanismo. Il plateau di Ontong Giava, la maggiore provincia magmatica, ha un'area di quasi due milioni di chilometri quadrati. Lo studio di queste enormi strutture sta chiarendo i meccanismi interni della Terra e i loro effetti sulle condizioni ambientali alla superficie.

le scienze della Terra e le risorse naturali di Hannover (Germania) ipotizzò, basandosi sull'analisi delle registrazioni di onde sismiche riflesse, l'esistenza di estese colate laviche stratificate in corrispondenza di molti margini continentali sommersi. In seguito vari ricercatori ottennero dati sismici più precisi che sembrano avvalorare la tesi di Hinz. Infatti rocce ignee recuperate in perforazioni eseguite sui margini continentali al largo dell'Irlanda e della Norvegia confermano l'origine vulcanica delle strutture individuabili attraverso i profili sismici a riflessione.

Un'ulteriore conferma viene da analoghi studi sismici eseguiti presso il plateau delle Kerguelen, che mostrano con evidenza come i plateau oceanici siano costituiti principalmente da rocce vulcaniche. Più di recente, esperti a bordo della *JOIDES Resolution* hanno ottenuto campioni di roccia da due enormi plateau oceanici: quello delle Kerguelen nel 1988 e quello di Ontong Giava, situato nell'Oceano Pacifico a nord-est dell'Australia, nel 1990. Questi campioni sono costituiti da basalti simili a quelli che vengono rinvenuti negli espandimenti basaltici continentali.

L'interrogativo che gli studiosi si sono posti riguarda le modalità di formazione di queste imponenti strutture vulcaniche. Sfortunatamente, però, le informazioni fondamentali necessarie per rispondere a questa domanda - cioè il volume totale della lava e delle rocce intrusive contenute nelle province magmatiche e la velocità con cui queste si sono formate - sono tuttora scarse. Nelle aree continentali si può misurare direttamente l'area degli espandimenti basaltici, ma si può solo ipotizzarne lo spessore. La crosta al di sotto dei continenti raggiunge in media i 35 chilometri di spessore, e soltanto la sua parte sommitale è accessibile all'esplorazione. Al di sotto degli oceani la situazione è ancora più complessa; infatti il campionamento anche solo della porzione più superficiale dei 5-25 chilometri di crosta presenti in corrispondenza dei plateau oceanici necessita dell'impiego di costose navi da perforazione.

Il trascorrere del tempo contribuisce inoltre a mascherare la reale estensione delle antiche strutture vulcaniche. Le più antiche province magmatiche sono spesso intensamente erose e, di conseguenza, alterate e ridotte di dimensioni. John J. Mahoney dell'Università di Hawaii valuta, per esempio, che i trappi del Deccan avrebbero originariamente coperto un'area tre volte maggiore dell'attuale. Inoltre, durante i 65 milioni di anni trascorsi dalla loro messa in posto, l'espansione dei fondi oceanici ha, a quanto pare, trasportato parte delle lave originarie fino alle Seychelles, al plateau delle Mascarene e alla Dorsale Chagos-Laccadive, situate tutte molto a sud e a sud-ovest dell'India.

Anche i plateau oceanici e i margini vulcanici sottomarini hanno in parte perduto il loro aspetto originario a causa dell'alterazione. I campioni ottenuti dalle perforazioni e le immagini sismiche delle province magmatiche del margine continentale norvegese e del plateau delle Kerguelen dimostrano che subito dopo la loro formazione queste strutture emergevano dal mare e solo in un secondo tempo sprofondarono gradualmente fino alla loro attuale posizione. Coffin ha stimato che i plateau delle Kerguelen e di Broken Ridge, come la maggior parte dei plateau e delle dorsali dell'Oceano Indiano, si formarono al di sopra del livello marino e vi rimasero successivamente anche per 50 milioni di anni, un periodo di tempo sufficiente a permettere la loro erosione.

Nonostante la scarsità di dati disponibili e le intense trasformazioni fisiche subite dalle grandi province magmatiche, siamo riusciti a dedurre l'area originariamente coperta dalla lava relativamente a cinque dei plateau meglio conosciuti: quelli di Ontong Giava e di Kerguelen-Broken Ridge, i margini vulcanici nordatlantici e gli espandimenti basaltici continentali del Deccan e del Columbia River. Questi ultimi, sebbene siano la meno estesa fra le province magmatiche considerate, comprendono una regione più vasta dell'Italia settentrionale; il più grande di tutti, il plateau di Ontong Giava, ha una dimensione pari a due terzi dell'Australia. Per poter stimare l'area e il volume delle grandi province magmatiche, i geologi hanno compreso la necessità di analizzare strutture attuali di tipo analogo.

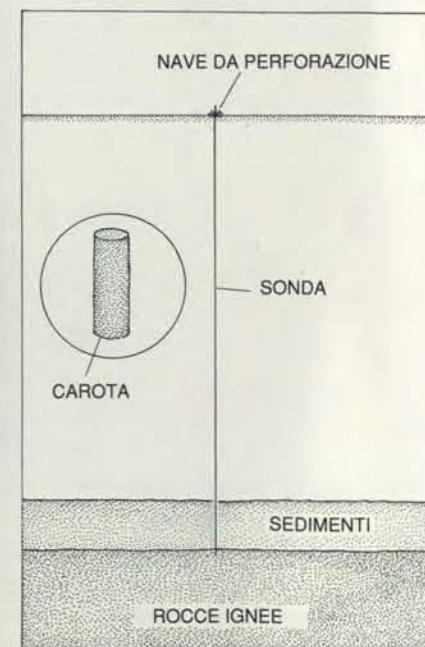
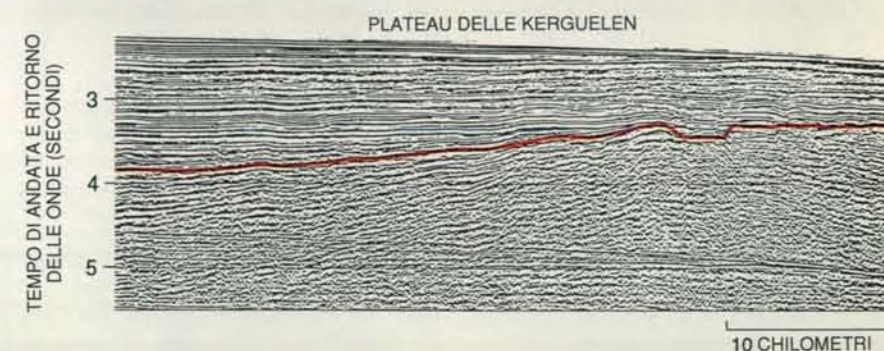
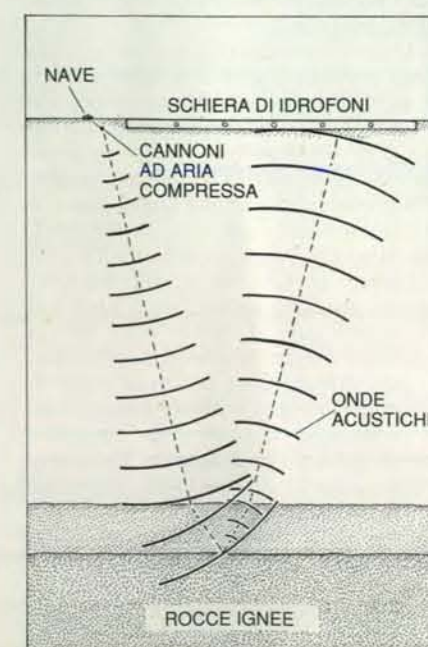
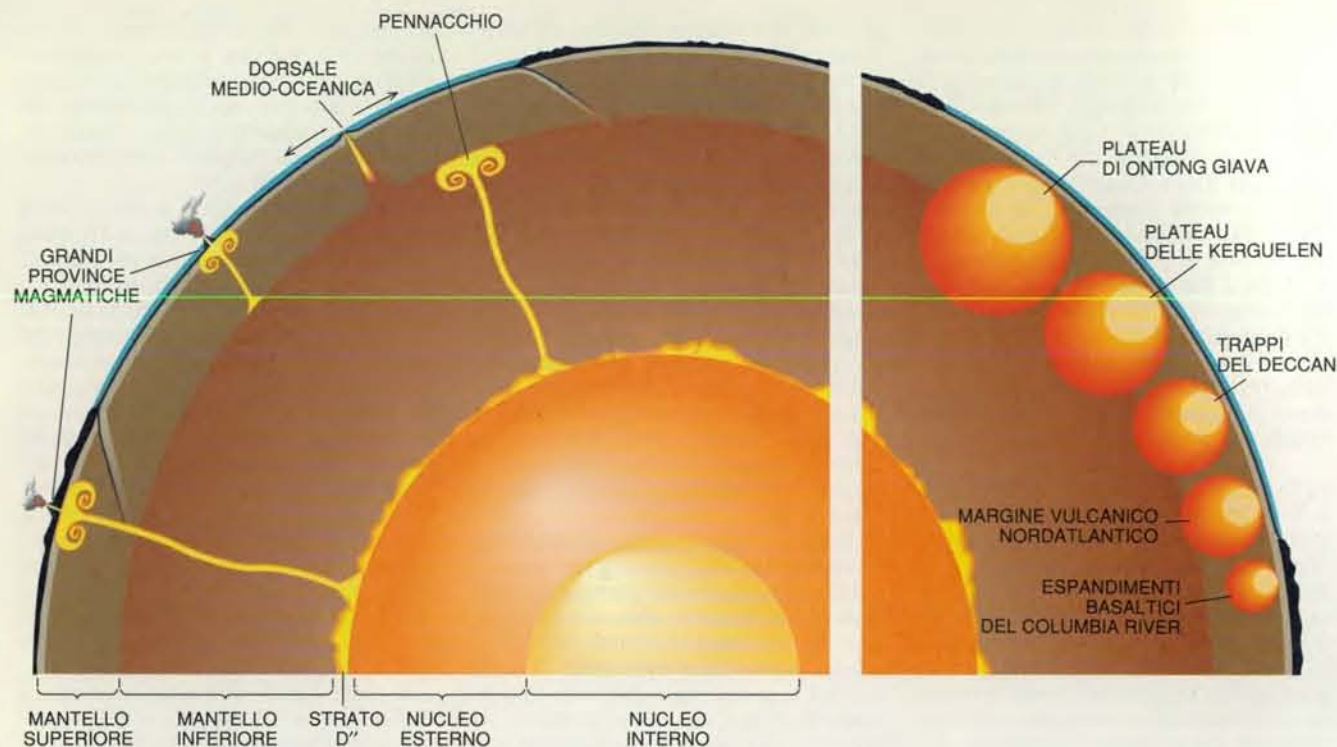
Nessuna delle regioni vulcaniche attualmente attive può essere anche lontanamente paragonata per dimensioni alle grandi province magmatiche estinte. Anche i punti caldi caratterizzati da maggiore attività (la grande isola di Hawaii e quella di Réunion, situata nell'Oceano Indiano a est del Madagascar) coprono un'area della crosta terrestre pari soltanto a circa un quinto delle dimensioni dei basalti del Columbia River. Tuttavia gli studi sulla struttura crostale sottostante ad Hawaii contribuiscono alla conoscenza dei meccanismi generali che stanno alla base di tutti i punti caldi.

Nel 1982 Anthony B. Watts e Uri S. ten Brink, assieme ai loro collaboratori del Lamont-Doherty Earth Observatory, raccolsero dati sismici registrando onde sonore generate artificialmente che vengono riflesse e rifratte al di sotto delle isole e dei vulcani sottomarini delle Hawaii. La velocità di queste onde dipende dalle caratteristiche fisiche della roccia attraversata (come la densità e le costanti elastiche); laddove questi parametri cambiano, anche le onde subiscono variazioni nel loro percorso. Le analisi sismiche rivelano che le Isole Hawaii si formarono quando il magma basaltico si intruse all'interno della preesistente crosta oceanica e si accumulò sopra di essa.

Tuttavia questi magmi superficiali evidentemente non contengono tutto il materiale associato alla formazione delle isole. Al di sotto di queste e dei vulcani sottomarini esiste una zona anomala di roccia, individuata dalla forte velocità con cui le onde longitudinali la attraversano. Si reputa che questa struttura sepolta sia stata prodotta dalla stessa sorgente del mantello che generò i rilievi. Le tecniche attuali non sono in grado di rivelare quanto nuovo materiale igneo si sia aggiunto alla crosta oceanica già esistente via via che le isole andavano formandosi. I valori calcolati del volume di materiale derivato da pennacchi del mantello e contenuto nella catena hawaiiana devono pertanto essere considerati valori minimi.

Siamo riusciti tuttavia a stabilire una relazione fondamentale tra la struttura dei dati di velocità sismica (e altri dati geofisici) e il volume totale del materiale compreso nell'ambito delle Isole Hawaii. Successivamente abbiamo utilizzato queste isole come modello per dedurre la distribuzione dei materiali vulcanici in diverse antiche e grandi province magmatiche. Secondo i nostri calcoli una piccola provincia come quella del Columbia River è costituita approssimativamente da 1,3 milioni di chilometri cubi di roccia, mentre la provincia gigante di Ontong Giava contiene almeno 36 milioni di chilometri cubi di roccia ignea, una quantità sufficiente per seppellire gli Stati Uniti sotto una coltre di basalto di quasi 5 chilometri di spessore.

Il passo successivo nella conoscenza dell'origine e del significato delle varie province magmatiche è quello di scoprire quanto velocemente si siano formate. Si accumularono in modo lento e costante lungo un periodo di decine di milioni

La nave da ricerca *JOIDES Resolution* è la protagonista dell'Ocean Drilling Program, il più ampio progetto internazionale di studio nel settore delle scienze della Terra. La nave trasporta una sonda in grado di penetrare nei sedimenti del fondo oceanico e di prelevare carote anche a una profondità di due chilometri. Grazie a questi sondaggi è possibile analizzare direttamente le rocce ignee provenienti dai margini continentali sommersi e dai plateau oceanici.

La nave *Rig Seismic* della Australian Geological Survey Organization raccoglie dati sismici sulla morfologia del fondo oceanico. Una schiera di cannoni ad aria compressa genera onde di tipo acustico che sono riflesse dal fondo (*qui sotto*). Idrofoni trainati dalla nave captano le onde riflesse che sono poi registrate. Una immagine sismica così ottenuta (*in basso a destra*) rivela la struttura magmatica del plateau delle Kerguelen (*area sottostante la curva in rosso*).

Pennacchi di materiale incandescente che risalgono attraverso il mantello danno origine alle grandi province magmatiche quando la loro estremità superiore raggiunge la superficie (*a sinistra*). Probabilmente i pennacchi hanno origine nello strato D'', posto tra il nucleo e il mantello, e nella regione intermedia tra mantello superiore e inferiore. Le colonne di roccia fusa da cui vennero generate le province magmatiche più massicce dovevano essere così enormi da far pensare che avessero origine almeno in parte nel mantello inferiore, molto probabilmente nello strato D''. Le sfere a destra rappresentano il diametro minimo (*in giallo*) e massimo (*in arancione*) dei pennacchi associati alle cinque maggiori province magmatiche.

di anni, con le medesime modalità con le quali si forma nuova crosta oceanica lungo le dorsali medio-oceaniche? Oppure nacquero da una serie di violentissime eruzioni vulcaniche che espulsero rapidamente gas e frammenti di roccia nell'aria e nell'acqua, trasformando bruscamente le caratteristiche geologiche di vastissime aree del globo?

Gli studiosi hanno fatto molti sforzi per tentare di rispondere a queste domande riguardo ai trappi del Deccan. Questo espandimento basaltico ha attirato particolare attenzione perché la sua formazione risale a circa 65 milioni di anni fa, proprio quando, alla fine del Cretaceo, si ebbe una delle maggiori estinzioni in massa, alla quale, anzi, potrebbe aver contribuito (si veda l'articolo *Che cosa causò l'estinzione in massa? Un'eruzione vulcanica* di Vincent E. Courtillot in «Le Scienze» n. 268, dicembre 1990). Robert A. Duncan della Oregon State University, insieme con Vincent E. Courtillot e Didier Vandamme dell'Institut de Physique du Globe di Parigi e con molti altri colleghi, ha effettuato una serie di datazioni con metodi radioattivi e di analisi magnetiche su campioni provenienti dai trappi del Deccan. I risultati ottenuti da questi ricercatori indicano che la maggior parte delle lave fu eruttata in un periodo di tempo inferiore a un milione di anni. Adoperando metodi simili, Ajoy K. Baksi della Louisiana State University stabilì che gli espandimenti basaltici del Columbia River furono effusi per la maggior parte in un milione e mezzo di anni.

Si possiedono pochi campioni in base ai quali stabilire direttamente l'età e la velocità di formazione dei plateau oceanici e dei margini vulcanici sommersi. Le scarse prove a disposizione fanno pensare per le province magmatiche sommerse a un accumulo altrettanto rapido di quello delle analoghe strutture continentali. Eldholm stima che la maggior parte della provincia vulcanica nordatlantica si sia formata in tre milioni di anni o anche meno. John A. Tarduno e collaboratori della Scripps Institution of Oceanography, in collaborazione con Mahoney, sono arrivati alla conclusione che il plateau di Ontong Giava si sia costituito in meno di tre milioni di anni. Le nostre analisi delle datazioni di rocce eseguite da Hugh Davies, allora alla Australian Geological Survey Organization, e da Hubert Whitechurch, allora all'Università di Strasburgo, mostrano che il plateau delle Kerguelen ebbe origine in massima parte in un arco di tempo di 4,5 milioni di anni.

Da un punto di vista geologico le più grandi province magmatiche si costituirono con eccezionale rapidità. Per fare un confronto, le Montagne Rocciose si sono innalzate in oltre 40 milioni di anni mentre la catena delle Isole Hawaii e i vulcani sottomarini della catena Emperor si sono formati in un arco di tempo di almeno 70 milioni di anni. Inoltre la maggior parte dell'attività vulcanica connessa alle province magmatiche potrebbe essersi verificata in episodi brevi ma violenti, intervallati da lunghi periodi di inattività.

Una volta raccolti i dati sul volume totale e sulla velocità di formazione delle maggiori province magmatiche, abbiamo potuto finalmente dedurre l'entità delle forze vulcaniche che le generarono. Le eruzioni che diedero origine alla provincia di Ontong Giava portarono all'emissione ogni anno di una quantità di roccia fusa compresa tra 12 e 15 chilometri cubi; il vulcanismo del Deccan ne produsse annualmente tra 2 e 8 chilometri cubi. Considerando che la formazione delle province magmatiche, al pari di ogni altro processo eruttivo, avviene a impulsi, in alcuni anni il ritmo di produzione crostale potrebbe anche essere stato molto maggiore. Per riferimento, Roger L. Larson dell'Università di Rhode Island calcola che il complesso globale delle dorsali medio-oceaniche ha prodotto annualmente una quantità compresa tra 16 e 26 chilometri cubi di crosta oceanica durante gli ultimi 150 milioni di anni. In altre parole quindi, le singole province magmatiche avrebbero generato nuova crosta a una velocità paragonabile, se non addirittura superiore, a quella che caratterizza attualmente il meccanismo globale di espansione dei fondi oceanici.

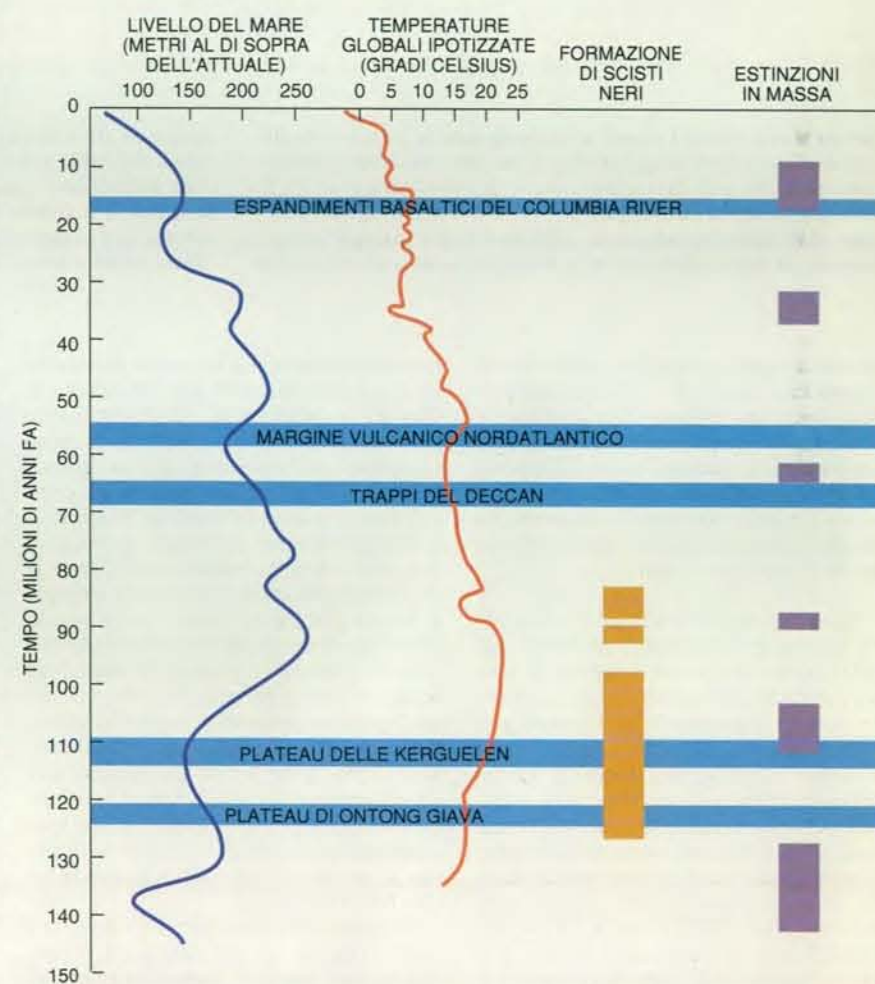
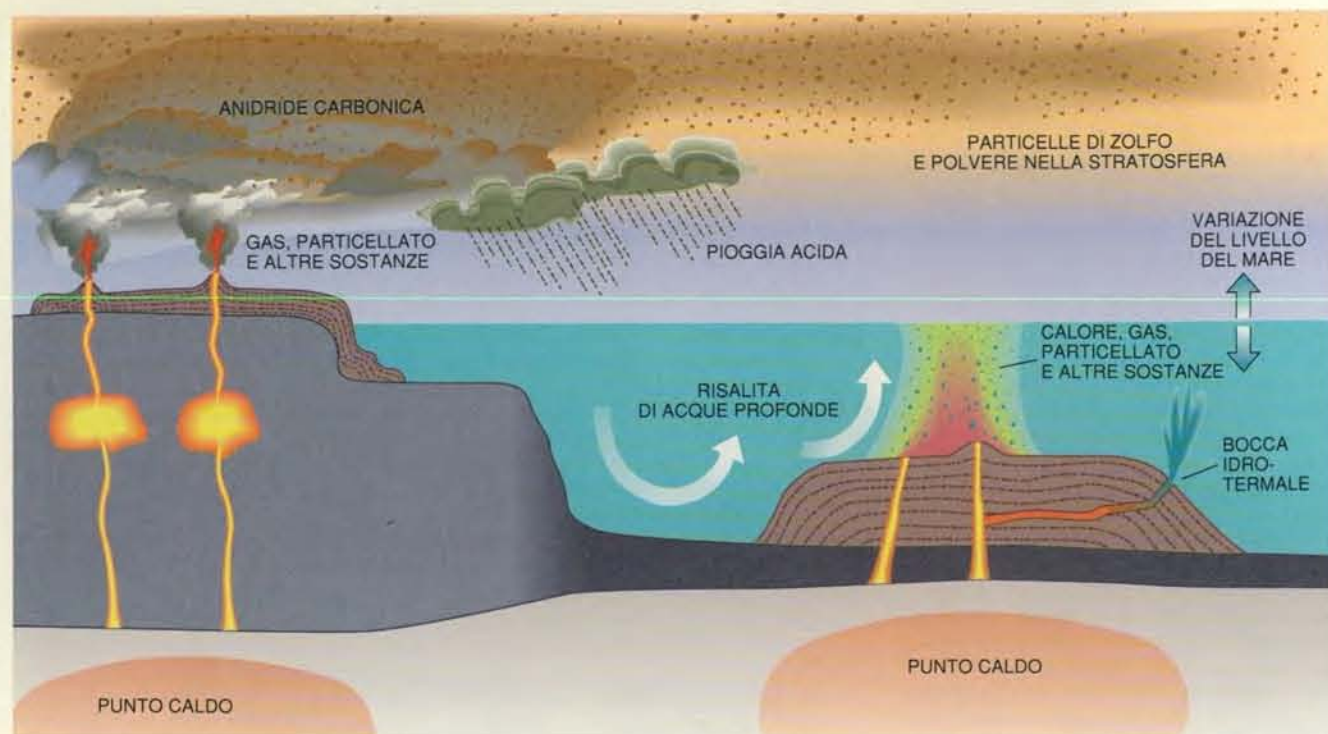
Dato che si sono formate così rapidamente, se raffrontate alla velocità dei sommovimenti interni della Terra, le grandi province magmatiche devono avere avuto origine da singole sorgenti isolate. Una volta noto il volume di roccia basaltica contenuta in queste province, si possono valutare le dimensioni dei pennacchi del mantello che le generano. Soltanto una parte del materiale che compone il pennacchio effettivamente fonde e raggiunge la crosta; la frazione di roccia fusa è probabilmente minore a grandi profondità, dove l'elevata pressione tende a mantenere il materiale del mantello allo stato solido. Di conseguenza, al di sotto della spessa litosfera continentale, la roccia dovrebbe fondere in misura minore rispetto a quanto avviene al di sotto della più sottile litosfera oceanica. Stephen M. Eggins della Australian National University e Shen-su Sun dell'Australian Geological Survey Organization stimano che la porzione fusa, basaltica del pennacchio (quella parte che provoca il vulcanismo alla superficie e le intrusioni magmatiche al di sotto di essa) rappresenti dal 5 al 30 per cento del suo volume totale.

Abbiamo usato questi valori per calcolare il diametro delle anomalie termiche del mantello che sono associate alle cinque province magmatiche meglio conosciute. Per semplificare abbiamo considerato che i pennacchi in risalita fossero sferici perché questa forma corrisponde al modo più efficiente in cui può avvenire il trasporto di materiale e permette al pennacchio stesso di muoversi a una velocità plausibilmente bassa nell'interno della Terra. Le nostre analisi indicano che la provincia di Ontong Giava deve avere avuto origine da un pennacchio con un diametro compreso fra un minimo di 600 chilometri e un massimo di 1400 chilometri.

Questo intervallo dimensionale è particolarmente significativo da un punto di vista geologico; indica infatti che i pennacchi più grandi devono contenere almeno un po' di materiale del mantello inferiore, che si trova a oltre 670 chilometri al di sotto della superficie. A quella profondità la velocità delle onde sismiche cambia all'improvviso, probabilmente in seguito a una variazione nella struttura mineralogica delle rocce che costituiscono il mantello. I geofisici continuano a discutere se l'intero mantello sia da considerarsi omogeneo oppure se il mantello inferiore e quello superiore si comportino come sistemi indipendenti con differenti meccanismi di circolazione.

Il nostro lavoro dà conferma ai modelli che prevedono almeno una certa interazione tra il mantello inferiore e quello superiore. Secondo noi, i pennacchi più grandi hanno origine nel mantello inferiore, molto probabilmente in corrispondenza dello strato D'', una regione con caratteristiche sismiche anomale situata appena all'esterno del nucleo. Pennacchi di dimensioni minori potrebbero invece risalire dalla zona di transizione tra mantello superiore e mantello inferiore, che si trova alla profondità di 670 chilometri.

Indipendentemente dal luogo di origine dei pennacchi del mantello, le loro caratteristiche e i loro effetti in superficie dipendono fortemente dalla temperatura, dalla composizione e dallo stato fisico delle rocce incontrate immediatamente al di sotto della litosfera. Questi fattori, unitamente alla resistenza locale della litosfera, determinano il volume, i tempi e i luoghi della fuoriuscita del magma alla superficie terrestre. I pennacchi che risalgono al di sotto delle masse continentali possono favorire la frammentazione dei continenti in corrispondenza di zone di debolezza strutturale, causando la formazione di estesi lineamenti vulcanici lungo i margini fratturati (si veda l'articolo *Vulcanismo di rift* di Robert S. White e Dan P. McKenzie in «Le Scienze» n. 253, settembre 1989). In certe circostanze il pennacchio può penetrare le spesse regioni centrali delle masse continentali, dando luogo a un espandimento basaltico.

Se invece il materiale del mantello in risalita emerge sul fondo marino, può creare un plateau oceanico. Esperimenti di laboratorio indicano che una lunga «coda» di materiale caldo dovrebbe seguire la testa sferica del pennacchio, dando luogo a una sorgente di magma duratura e di limitata estensione. Nel corso di milioni di anni il moto delle zolle provoca la migrazione del fondo ocea-

Fra gli effetti ambientali conseguenti alla formazione delle grandi province magmatiche vi sono oscillazioni del livello oceanico globale. Una massiccia attività vulcanica potrebbe alterare il chimismo e la circolazione sia dell'atmosfera sia degli oceani, influenzando l'evoluzione della vita. L'emissione di anidride carbonica da parte dei vulcani può contribuire all'effetto serra; il conseguente riscaldamento, accrescendo la produzione mondiale di biomassa, favorisce la formazione di scisti neri. Molte estinzioni in massa avvenute negli ultimi 150 milioni di anni coincidono con la formazione di province magmatiche, il che indica un rapporto di causa ed effetto.

Vari processi chimici e fisici accompagnano la formazione delle grandi province magmatiche. I vulcani emettono anidride carbonica che può far aumentare le temperature globali. La presenza nella stratosfera di polveri e particelle di zolfo prodotte dall'attività vulcanica può provocare piogge acide e bloccare la radiazione solare. Anche l'emissione di metalli pesanti e di altre sostanze potrebbe alterare la composizione del suolo, dell'aria e dell'acqua. Le variazioni topografiche associate a un esteso vulcanismo modificherebbero la circolazione oceanica e il livello del mare; la risalita di acqua profonda dovuta al riscaldamento e all'attività idrotermale trasformerebbe ulteriormente le condizioni dell'ambiente sottomarino.

nico al di sopra del punto caldo. Fuoriuscendo in superficie, la lava forma gradualmente una dorsale sottomarina lineare oppure una serie di isole e vulcani sottomarini. La catena Hawaii-Emperor si è sviluppata probabilmente in questo modo, sebbene rimanga un mistero come un pennacchio possa persistere per più di 70 milioni di anni.

Come dimostra l'eruzione del Pinatubo nelle Filippine, avvenuta nel 1991, anche eruzioni vulcaniche di portata relativamente modesta possono deteriorare gravemente le condizioni ambientali locali. Tuttavia questi eventi geologici rappresentano episodi trascurabili rispetto agli enormi sconvolgimenti che avvengono durante la formazione delle grandi province magmatiche. Si può quindi intuire come quelle antiche eruzioni abbiano avuto conseguenze profonde. Nel 1972 Peter R. Vogt del Naval Research Laboratory propose per primo l'ipotesi che l'arrivo in superficie di un pennacchio del mantello possa provocare cambiamenti fisici e chimici a livello globale; gli effetti ambientali associati a questi cambiamenti potrebbero avere effetti rilevanti sull'evoluzione della vita.

La recente attenzione per la teoria dell'effetto serra e dei cambiamenti climatici a livello globale ha prodotto un rinnovato interesse per le ipotesi formulate da Vogt. Stephen Self dell'Università di Hawaii e Michael R. Rampino della New York University sottolineano come l'impatto sull'ambiente di una grande provincia magmatica dipenda in parte dal fatto che essa si formi sulla terraferma o nell'oceano. I plateau oceanici e i margini vulcanici passivi modificano la geometria dei bacini oceanici e alterano il livello globale del mare. Noi stimiamo che l'accumulo di magma che diede origine al plateau di Ontong Giava abbia alzato il livello del mare di circa 10 metri. I plateau oceanici e i margini vulcanici possono ostacolare o incanalare i movimenti delle acque in maniera tale da influenzare la circolazione, l'erosione e la sedimentazione a grande scala, specialmente se l'attività vulcanica si imposta in un sito di equilibrio instabile tra due bacini oceanici.

A causa della loro maggiore densità e massa rispetto all'atmosfera, gli oceani hanno una capacità molto maggiore di assorbire e diluire i gas e il calore; questa caratteristica tende a far sì che le conseguenze delle eruzioni vulcaniche sottomarine risultino attenuate. D'altro canto il vulcanismo sottomarino e l'attività idrotermale a esso associata possono liberare metalli in traccia, come per esempio l'arsenico, che sono tossici per gli organismi marini. Il calore prodotto da queste eruzioni determina la risalita delle acque profonde, alterando la circolazione di quelle superficiali e perturbando l'equilibrio degli organismi che vi abitano.

L'anidride carbonica emessa nei pressi di un sito vulcanico sottomarino si diffonde nell'oceano e, così facendo, può aumentare l'alcalinità dell'acqua e influire sia sulla vita marina sia sul clima. Se livelli elevati di anidride carbonica dovessero portare a un riscaldamento globale, la circolazione oceanica rallenterebbe e trasporterebbe quindi una minor quantità di ossigeno disciolto. Come è stato documentato dal compianto Seymour O. Schlanger della Northwestern University, da Michael A. Arthur della Pennsylvania State University e da Hugh C. Jenkyns dell'Università di Oxford, le acque impoverite di ossigeno possono contribuire alla formazione di scisti neri ricchi di carbonio.

L'atmosfera deve avere assorbito una enorme quantità di anidride carbonica intorno a 120 milioni di anni fa, quando l'attività vulcanica raggiunse il suo culmine. Larson ipotizza che una rilevante instabilità termica nello strato D″ abbia dato origine a «superpennacchi» che infine generarono i plateau oceanici di Ontong Giava e delle Kerguelen, assieme a molti altri più piccoli nell'Oceano Pacifico. Kenneth G. Caldeira della Pennsylvania State University e Rampino hanno elaborato modelli al calcolatore del ciclo geochimico basandosi sulla quantità ipotizzata di anidride carbonica presente nell'atmosfera a quell'epoca. A partire dai modelli, essi hanno supposto che le temperature globali fossero in media da 7,6 a 12,5 gradi Celsius più alte di quelle attuali, tenendo conto della configurazione geografica differente e del livello del mare più alto.

Le eruzioni vulcaniche che avvengono su terraferma alterano direttamente le caratteristiche chimiche e fisiche dell'atmosfera. Alan R. Huffman della Exxon Exploration Company di Houston calcola che un singolo evento vulcanico che generi un espandimento basaltico di 1000 chilometri cubi (il volume di una tipica effusione lavica nella provincia del Columbia River) emette 16 miliardi di tonnellate di anidride carbonica, 3 miliardi di tonnellate di zolfo e 30 milioni di tonnellate di alogeni. Le migliaia di episodi di questa entità che dovettero avvenire per generare una singola grande provincia magmatica avrebbero modificato a tal punto l'atmosfera da far apparire al confronto insignificanti gli effetti dell'inquinamento attuale prodotto dall'uomo.

Secondo Rampino e Self le eruzioni esplosive di lave ricche in silice spesso trasportano particelle di zolfo nella stratosfera, all'interno della quale vengono trasformate in minuscole goccioline di acido solforico. Le lave basaltiche producono una quantità di zolfo 10 volte maggiore per unità di volume rispetto a quelle ricche in silice; Charles B. Officer e collaboratori del Dartmouth College hanno calcolato che se i gas e le particelle prodotti durante un'eruzione basaltica fossero scagliati nella stratosfera, potrebbero causare a breve scadenza piogge acide, nonché oscuramento e raffreddamento a livello planetario. Richard B. Stothers del Goddard Space Flight Center della NASA e molti altri ipotizzano che le enormi fontane di lava e l'energica convezione atmosferica al di sopra delle bocche eruttive negli espandimenti basaltici continentali possano fornire una spinta verso l'alto sufficiente per immettere materiale vulcanico nella stratosfera.

La potentissima eruzione del Lakagigar, avvenuta in Islanda a cavallo tra il 1783 e il 1784, illustra gli effetti potenzialmente distruttivi del vulcanismo basaltico. Sebbene essa abbia prodotto solo 15 chilometri cubi di lava, le temperature locali diminuirono notevolmente durante gli anni successivi. Circa tre quarti del bestiame dell'Islanda perì in seguito sia al deterioramento del clima sia all'emissione di gas acidi; la carestia che ne derivò condusse a morte circa un quarto dell'intera popolazione islandese. Veli di polvere, nebbie e foschie apparvero su gran parte dell'Europa e sulle regioni adiacenti dell'Asia e dell'Africa per molti mesi dopo l'eruzione.

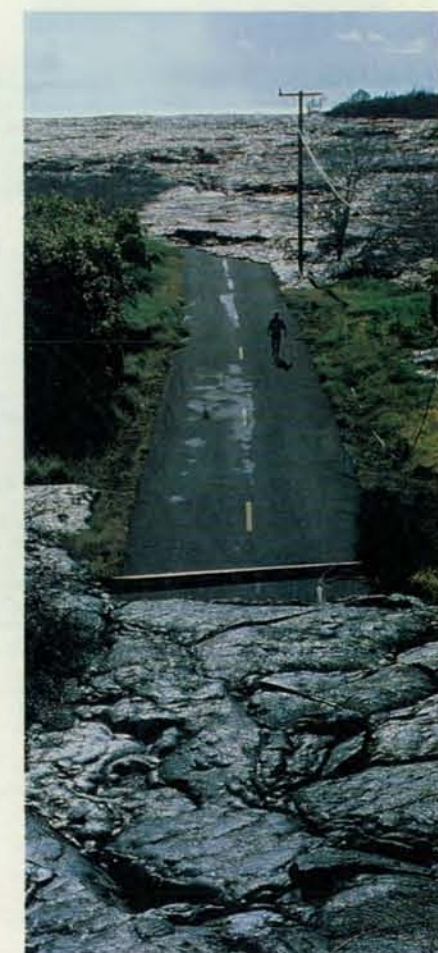
Si stanno scoprendo prove sempre più convincenti del fatto che le conseguenze ambientali del vulcanismo basaltico abbiano contribuito effettivamente alle estinzioni in massa. La più grande estinzione della storia terrestre avvenne 248 milioni di anni fa, quando si formarono i trappi della Siberia. A quell'epoca scomparve circa il 95 per cento di tutte le specie marine; la conseguente disponibilità di spazi evolutivi favorì la comparsa sulla scena dei dinosauri.

Le ripercussioni biologiche della formazione delle grandi province magmatiche possono in parte dipendere dallo stato dell'ambiente globale in quel dato momento. Quando l'equilibrio ambientale è già sollecitato da altri fattori, questi episodi di vulcanismo possono scatenare rapidi cambiamenti del clima, degli oceani e della biosfera; se, viceversa, le condizioni ambientali sono stabili, nella documentazione geologica si riscontrano scarsi effetti.

Stranamente, la provincia magmatica più imponente, quella di Ontong Giava, non produsse a quanto pare alcuna estinzione; al contrario, la sua origine coincide con la deposizione di scisti neri, indicativa di una fase di intensificata attività biologica. L'ipotesi da noi proposta è che gli effetti deleteri associati alla formazione di questo plateau siano stati minimi, in parte perché l'eruzione fu sottomarina, come indicano gli studi dell'Ocean Drilling Program condotti da Loren Kroenke dell'Università di Hawaii e Wolfgang H. Berger della Scripps Institution of Oceanography. Inoltre noi riteniamo che 120 milioni di anni fa l'ambiente del nostro pianeta fosse in una condizione più elastica dell'attuale, sebbene un'asserzione di questo tipo sia difficile da provare e quantificare. La formazione dei meno estesi plateau delle Kerguelen e di Broken Ridge, avvenuta circa 110 milioni di anni fa, coincise invece con un'estinzione in massa oltre che con una rilevante deposizione di scisti neri.

Le eruzioni che diedero luogo ai trappi del Deccan e al vulcanismo coevo lungo i margini dell'India occidentale e lungo i plateau delle Seychelles e delle Mascarene avvennero approssimativamente 65 milioni di anni fa, proprio all'epoca in cui si ebbe l'estinzione di quasi la metà di tutte le specie viventi, compresi i dinosauri. Si continua a discutere vivacemente se questa estinzione sia da attribuire all'impatto di un grande asteroide o a cause di natura più prettamente terrestre. Anche se fosse vera la teoria dell'asteroide, le eruzioni del Deccan potrebbero comunque avere contribuito a un deterioramento ambientale tale da amplificare le ripercussioni dell'impatto.

Un altro cambiamento ambientale significativo, benché meno conosciuto, ebbe luogo 10 milioni di anni dopo la fine dei dinosauri, durante la formazione dei margini vulcanici nordatlantici. In quel periodo si verificò un'intensa attività idrotermale e scomparvero molti foraminiferi bentonici e mammiferi terrestri. David K. Rea e collaboratori dell'Università del Michigan, assieme a Ellen Thomas della Yale University, hanno trovato prove di grandi variazioni sia nella circolazione delle acque profonde sia in quella atmosferica.

L'analisi degli isotopi dell'ossigeno nei gusci dei foraminiferi indica che durante l'Eocene inferiore, circa 55 milioni di anni fa, la temperatura degli oceani era più alta di quanto fosse mai stata nei precedenti 70 milioni d'anni. Eldholm, lavorando con la Thomas, ha supposto recentemente che la temperatura mite degli oceani dell'Eocene inferiore fosse dovuta all'emissione di anidride carbonica in seguito alle eruzioni lungo i mar-

Le colate laviche del Kilauea ad Hawaii sono solo un pallido esempio dell'ampiezza e del potenziale impatto ambientale delle effusioni di materiale incandescente connesse con le grandi province magmatiche. Singoli strati di lava negli espandimenti basaltici del Columbia River, per esempio, si estendono per centinaia di chilometri e costituiscono depositi di parecchi metri di spessore.

gini vulcanici nordatlantici. Strati di ceneri databili a 55 milioni di anni fa sono presenti in vaste regioni dell'Europa nordoccidentale, a conferma della violenza di quell'episodio di vulcanismo. L'aumento della temperatura atmosferica potrebbe aver condotto alla formazione di uno strato di acqua relativamente calda alla superficie degli oceani alle alte latitudini; la presenza di questo strato caldo avrebbe impedito il mescolamento con le acque sottostanti, più fredde, e la conseguente variazione nella circolazione delle acque profonde potrebbe essere stata fatale per molte specie bentoniche.

Anche gli espandimenti basaltici del Columbia River, relativamente meno estesi, coincidono con un'estinzione in massa avvenuta 16 milioni di anni fa. Come viene puntualizzato da Maureen E. Raymo del Massachusetts Institute of Technology, proprio a quell'epoca iniziò a instaurarsi il ciclo delle glaciazioni; è probabile che l'equilibrio ambientale globale fosse già così delicato che anche un'eruzione di modesta entità poteva avere su di esso conseguenze di notevole rilevanza.

Come è ovvio, i geologi hanno appena iniziato a scalfire - in senso sia letterale sia figurato - la superficie delle grandi province magmatiche. Le attuali conoscenze su queste strutture dimostrano ampiamente che esse contengono informazioni di importanza cruciale sui meccanismi interni della Terra e sulle cause naturali dei cambiamenti a scala globale. Attualmente si stanno intensificando gli sforzi per ottenere profili sismici con risoluzione migliore, intraprendere ulteriori studi, sia sul campo sia in laboratorio, costruire modelli più perfezionati e condurre perforazioni più sistematiche. Questi lavori consentiranno senza dubbio di comprendere meglio gli stretti legami tra l'interno e l'esterno della Terra.

BIBLIOGRAFIA

MACDOUGALL J. D. (a cura), *Continental Flood Basalts*, Kluwer Academic Publishers, 1988.

DUNCAN R. A. e RICHARDS M. A., *Hotspots, Mantle Plumes, Flood Basalts, and True Polar Wander* in «Reviews of Geophysics», 29, n. 1, febbraio 1991.

LARSON ROGER L., *Geological Consequences of Superplumes* in «Geology», 19, n. 10, ottobre 1991.

STOREY B. C., ALABASTER T. e PANKHURST R. J. (a cura), *Magmatism and the Causes of Continental Break-up*, Special Publication n. 68, Geological Society of London, 1992.

COFFIN M. F. ed ELDHOLM O., *Large Igneous Provinces: Crustal Structure, Dimensions, and External Consequences* in «Reviews of Geophysics» (in stampa).

Le basi molecolari del tetano e del botulismo

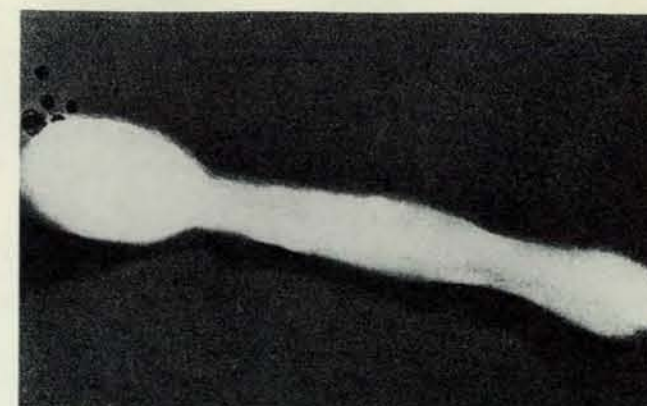
Le neurotossine responsabili di queste due sindromi sono zincoproteasi neurospecifiche che, penetrando all'interno delle cellule nervose, bloccano il rilascio dei neurotrasmettitori e causano le tipiche paralisi

di Giampietro Schiavo, Ornella Rossetto e Cesare Montecucco

Il tetano e il botulismo sono sindromi neuroparalitiche dovute all'intossicazione con neurotossine proteiche prodotte da batteri anaerobi.

A causa dei suoi sintomi eclatanti e caratteristici, il tetano è noto sino dagli albori della medicina e la parola stessa deriva dal greco τέτανος (contrattura). La malattia, descritta per la prima volta da Ippocrate ventiquattro secoli fa, si manifesta con una spasmodica contrattura muscolare generalizzata (paralisi spastica) che inizia dai muscoli facciali e del collo per estendersi poi ai muscoli della colonna vertebrale, dell'addome e degli arti. La contrattura dei muscoli masseteri, con impedimento dell'apertura delle mascelle, e dei muscoli mimici sono all'origine del *trismus* e del *risus sardonicus*, due aspetti caratteristici del tetano. Il paziente tetanizzato mostra dispnea, copiosa sudorazione, ipertermia. Anche con l'impiego delle moderne tecniche di rianimazione, la malattia ha spesso un esito fatale per collasso cardiaco e insufficienza respiratoria; nel caso che il paziente riesca a superare la fase acuta rimangono spesso danni neurologici permanenti. Raramente il tetano può manifestarsi anche in forma locale, e in questo caso la paralisi spastica è limitata a una parte dell'organismo.

A causa della minore evidenza dei sintomi, il botulismo è passato per molto tempo inosservato. Questo termine che deriva da *botulus* (salsiccia, in latino) è apparso per la prima volta nella letteratura medica solo alla fine del XVIII secolo in seguito a un caso di botulismo verificatosi in Germania: tredici persone avevano condiviso delle salsicce e sei ne erano morte. Al contrario del tetanizzato, il paziente affetto da botulismo mostra una paralisi flaccida, caratterizzata da debolezza muscolare e disturbi oculari con diplopia, cioè incapacità di mettere a fuoco un oggetto per sfalsamento dei piani visivi, e perdita del riflesso alla luce. Si notano poi difficoltà del movimento e della deglutizione per debolezza, perdita di coordinazione dei muscoli faringei e cessazione della salivazione. Seguono la perdita di controllo degli sfinteri e dei muscoli volontari mentre il paziente rimane perfettamente cosciente e vigile. La mortalità per botulismo, sebbene elevata, è inferiore a quella per tetano e interviene principalmente per paralisi dei muscoli respiratori; nei sopravvissuti le manifestazioni cliniche scompaiono dopo alcune settimane.

Per molti secoli si è creduto che il tetano fosse una sindrome di natura neurologica fino a che, nel 1884, Antonio Carle e Giorgio Rattone, dell'Istituto di patologia generale dell'Università di Torino, dimostrarono l'eziologia infettiva della malattia, inducendo il tetano in conigli inoculati con materiale prelevato da una pustola di un uomo deceduto per tetano. Subito dopo, il medico tedesco Arthur Nicolaier identificò il batterio come un bacillo allungato, spesso con spora in posizione subterminale. Nel 1889 Shibasaburo Kitasato, all'Istituto di igiene dell'Università di Berlino, mise a punto una procedura, utilizzata in seguito innumerevoli volte in microbiologia, per ottenere colture pure di un batterio sporulante. Riscaldando a 80 gradi Celsius brodo seminato con pus tetanico, riuscì a uccidere i batteri, ma non le spore.

Allo stesso tempo Guido Tizzoni e Giuseppina Cattani dell'Istituto di patologia generale dell'Università di Bologna arrivarono al medesimo risultato utilizzando però una tecnica di passaggi successivi su piastra di agar. Il nuovo batterio fu denominato *Clostridium tetani* per la somiglianza dei batteri sporigeni a una clava (in latino *clostridium*). Questo risultato fu di fondamentale importanza per il passo successivo: la scoperta che il tetano è causato da una tossina, così come Emile Roux e Alexandre Yersin avevano dimostrato per la difterite nel 1888 all'Istituto Pasteur di Parigi. Nei filtrati acellulari di colture di *Cl. tetani* Tizzoni e Cattani a Bologna e Knut Faber in Danimarca riuscirono a dimostrare nel 1890 la presenza di una tossina proteica, la tossina tetanica, sola responsabile di tutti i sintomi clinici del tetano. Fu così dimostrato che il tetano è una intossicazione e non un'infezione.

Sulla base di questo stesso percorso sperimentale, Emile van Ermengem in Belgio, analizzando un caso di botulismo che coinvolse numerose persone, le quali avevano ingerito del prosciutto affumicato, isolò dalla carne un nuovo bacillo, cui diede il nome di *Bacillus botulinum*. Egli dimostrò inoltre che anche nel botulismo la sola causa della malattia era una tossina proteica, prodotta dal batterio coltivato in condizioni anaerobie. Il batterio fu successivamente classificato nel genere *Clostridium* e nel tempo sono stati identificati sette tipi immunologicamente distinguibili di tossina botulinica denominati con le lettere da A a G. I sierotipi A, B ed E (rarissimamente F) sono coinvolti nel botulismo umano, mentre i sierotipi C e D sono associati quasi esclusivamente al botulismo animale.

Queste tossine sono i più potenti veleni sinora noti: iniettate intramuscolarmente ne bastano pochi miliardesimi di grammo per uccidere un uomo e quindi non c'è da stupirsi che esse siano state considerate come possibili armi biologiche. Ciò appare tuttavia poco probabile alla luce del fatto che la loro tossicità per via orale o naso-faringea è molto più bassa.

Le spore di *Cl. tetani* e di *Cl. botulinum* sono ubiquitarie e resistenti all'essiccazione: ne sono particolarmente ricchi i materiali biologici in decomposizione e le deiezioni di alcuni animali. La contaminazione di ferite e cibi con queste spore è quindi frequente. Anche piccole ferite poco profonde come un'abrasione, la foratura del lobo d'un orecchio, un tatuaggio, una puntura d'ago o di spina di rosa, possono permettere la penetrazione delle spore e creare condizioni favorevoli alla germinazione di *Cl. tetani* (ma molto raramente di *Cl. botulinum*) e alla produzione della tossina da parte dei batteri neoformati. A differenza di molti altri batteri, *Cl. tetani* non causa lesioni visibili, né invade i tessuti circostanti, e l'avvenuta infezione, e intossicazione, si manifesta, con i sintomi che abbiamo visto, anche due-quattro settimane dopo.

Una gravissima e frequente forma di contaminazione è quella che si verifica nei paesi meno sviluppati durante la recisione del cordone ombelicale con strumenti chirurgici non sterilizzati o comunque con metodi che non garantiscono l'asetticità. In alcune regioni africane si usa addirittura applicare, con presunte funzioni cauterizzanti, sterco animale, una sorgente ricchissima di spore. Secondo stime recenti dell'Organizzazione mondiale della sanità, il *tetanus neonatorum*, che frequentemente ne consegue, è responsabile della morte di almeno 450 000 neonati per anno ed è quindi nel mondo la singola maggior causa di morte nel primo anno di vita. Secondo la stessa fonte, gli adulti morti per tetano sono almeno 350 000 per anno, quasi tutti concentrati nelle zone più povere della Terra. Infatti, nel mondo sviluppato il tetano è stato ridotto a poche centinaia di casi grazie al ricorso alla vaccinazione.

Dato che la malattia è provocata da una tossina proteica, a Gaston Ramon, dell'Istituto Pasteur a Garches, all'inizio degli anni venti è bastato trattare la tossina con formaldeide per ottenere un tossoide tetanico, cioè un derivato privo di tossicità, ma ancora capace di stimolare il sistema immunitario. Ramon fu il primo a dimostrare, inoculandolo ad animali e a se stesso, che il tossoide tetanico induce uno stato di immunità attiva. L'efficacia del vaccino fu poi chiara a tutti quando si seppe dei 35 morti per tetano fra le truppe britanniche vaccinate impiegate nella seconda guerra mondiale a fronte del gran numero di casi verificatisi fra i soldati giapponesi e tedeschi, non vaccinati.

Il vaccino attuale è costituito dal tossoide adsorbito su ossido o fosfato di alluminio o fosfato di calcio, con funzioni di adiuvanti della risposta immunitaria. Il tossoide tetanico viene generalmente associato al tossoide difterico e questo è il prodotto più diffuso dell'industria biotecnologica mondiale. Si può stimare infatti che ne vengano prodotte circa 500 milioni di dosi per anno. Seguendo il protocollo di vaccinazione previsto dalla legislazione italiana si ottiene un titolo sierico di anticorpi antitossina tetanica e

L'incisione del neurologo scozzese Charles Bell rappresenta un paziente colpito da tetano in seguito a una ferita riportata nella battaglia di La Coruña (15 gennaio 1809) fra inglesi e francesi (Da *Anatomy and Philosophy of Expression*, 1832).

Queste microfotografie elettroniche di *Clostridium tetani* sono state ottenute da Franco Scanga con il primo microscopio elettronico in funzione in Italia all'Istituto superiore di sanità di Roma. La microfotografia a sinistra (a 20 000 ingrandimenti) mostra un batterio presente in una coltura ai primi stadi: sono molto evidenti i numerosi flagelli peritrichi. Nella microfotografia a destra (a 14 000 ingrandimenti) è particolarmente messo in rilievo in un batterio tipico di una coltura invecchiata l'ingrossamento della spora subterminale al quale è dovuta la caratteristica forma a clava da cui deriva il nome del batterio.

La cartina illustra la distribuzione del tetano neonatale nel mondo secondo i dati dell'Organizzazione mondiale della sanità relativamente al 1991. L'incidenza del tetano nell'adulto non è molto diversa; presenta però alti valori anche in Brasile.

antitossina difterica sufficiente a proteggere la maggior parte dei vaccinati per almeno una quarantina d'anni. I pochi casi di tetano che si registrano ancora in Italia riguardano anziani che non sono stati vaccinati o il cui livello di anticorpi antitossina tetanica sia sceso nel tempo al di sotto del livello minimo protettivo. In quest'ultimo caso basta una singola iniezione per riportare il titolo a valori sicuramente protettivi.

Il rischio dello sviluppo del tetano in un soggetto con ferite contaminate da spore tetaniche può essere prevenuto in modo passivo, mediante inoculazione di immunoglobuline prelevate da un soggetto immune, purché l'iniezione venga praticata molto presto dopo il trauma, in modo da prevenire l'entrata della tossina nelle cellule nervose. Se ciò avviene infatti gli anticorpi antitossina tetanica non hanno più alcun effetto. La virtuale scomparsa del tetano nel mondo occidentale è una delle grandi vittorie della medicina moderna derivante da ricerca di laboratorio, prima, e da un organizzato sistema di medicina sociale, poi.

Purtroppo, la situazione nelle regioni più arretrate del globo è molto differente. Per motivi economici e per la difficoltà di completare il ciclo di iniezioni previste (almeno tre intervallate di 2-4 settimane e 6-12 mesi), la maggioranza della popolazione del Terzo Mondo non possiede un livello di anticorpi antitossina tetanica sufficiente a garantire una efficace protezione e, di conseguenza, le madri non possono trasmettere ai nascituri i propri anticorpi protettivi. Per eliminare questa che è stata definita una «malattia senza scuse» l'Organizzazione mondiale della sanità è attualmente impegnata in un massiccio programma di immunizzazione che ha portato alla fine del 1991 a una copertura di oltre il 40 per cento della popolazione mondiale e che nell'ultimo decennio ha ridotto l'incidenza del tetano neonatale a meno della metà.

Molto rari sono nel mondo occidentale anche i casi di botulismo, ma non come conseguenza della vaccinazione, come avviene per il tetano, bensì in relazione alla diversa via di penetrazione nell'organismo delle tossine botuliniche. Infatti, il botulismo da ferita è molto raro. Il *Cl. botulinum* prolifera bene su carni o vegetali tenuti in anaerobiosi anche a basse temperature. In queste condizioni, i ceppi che portano un gene codificante per la tossina producono quest'ultima e la liberano complessata con altre proteine che la proteggono dall'acidità e dagli enzimi digestivi presenti nello stomaco. Nonostante ciò la dose di tossina ingerita viene in grandissima parte distrutta e questo spiega i molto più limitati effetti della somministrazione per via orale rispetto all'inoculazione intramuscolare. Il complesso botulinico che raggiunge indenne l'intestino, a causa del *p*H leggermente basico, si dissocia liberando la tossina che, se non viene digerita dagli enzimi pancreatici, può penetrare nell'organismo attraverso la parete intestinale. Il botulismo è ormai quasi scomparso nelle regioni più progredite del mondo per le migliorate norme di preparazione e conservazione dei cibi, ma permane come problema sanitario nelle zone più povere.

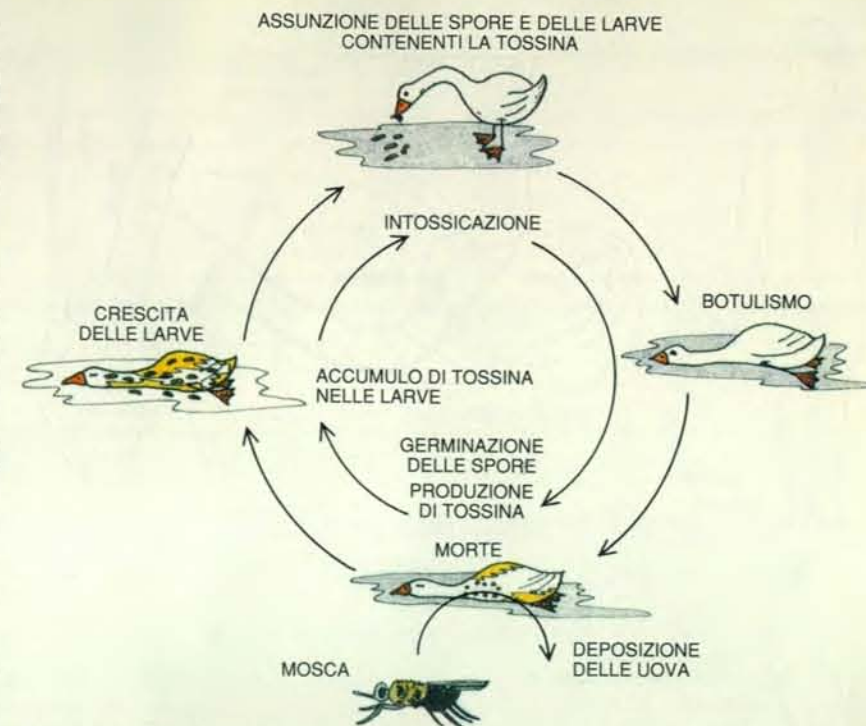
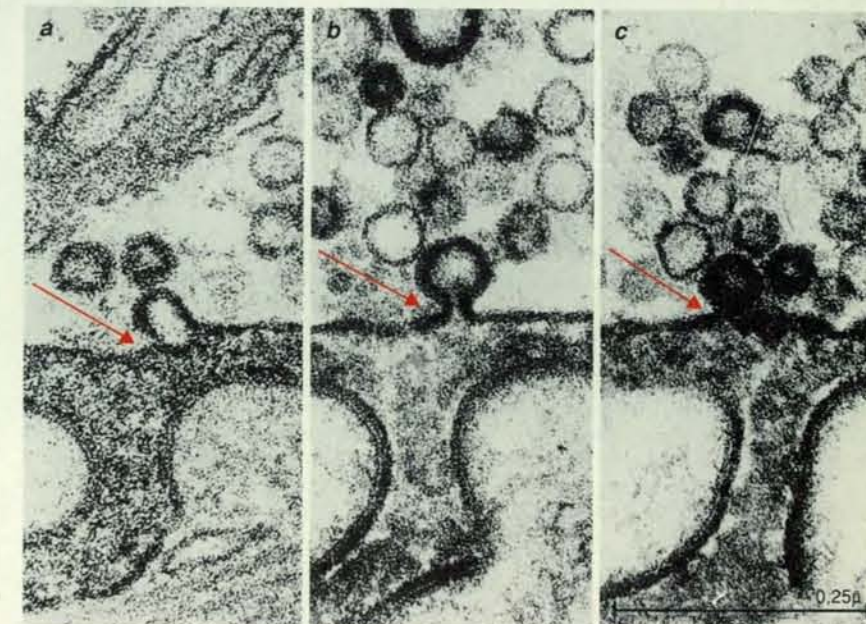
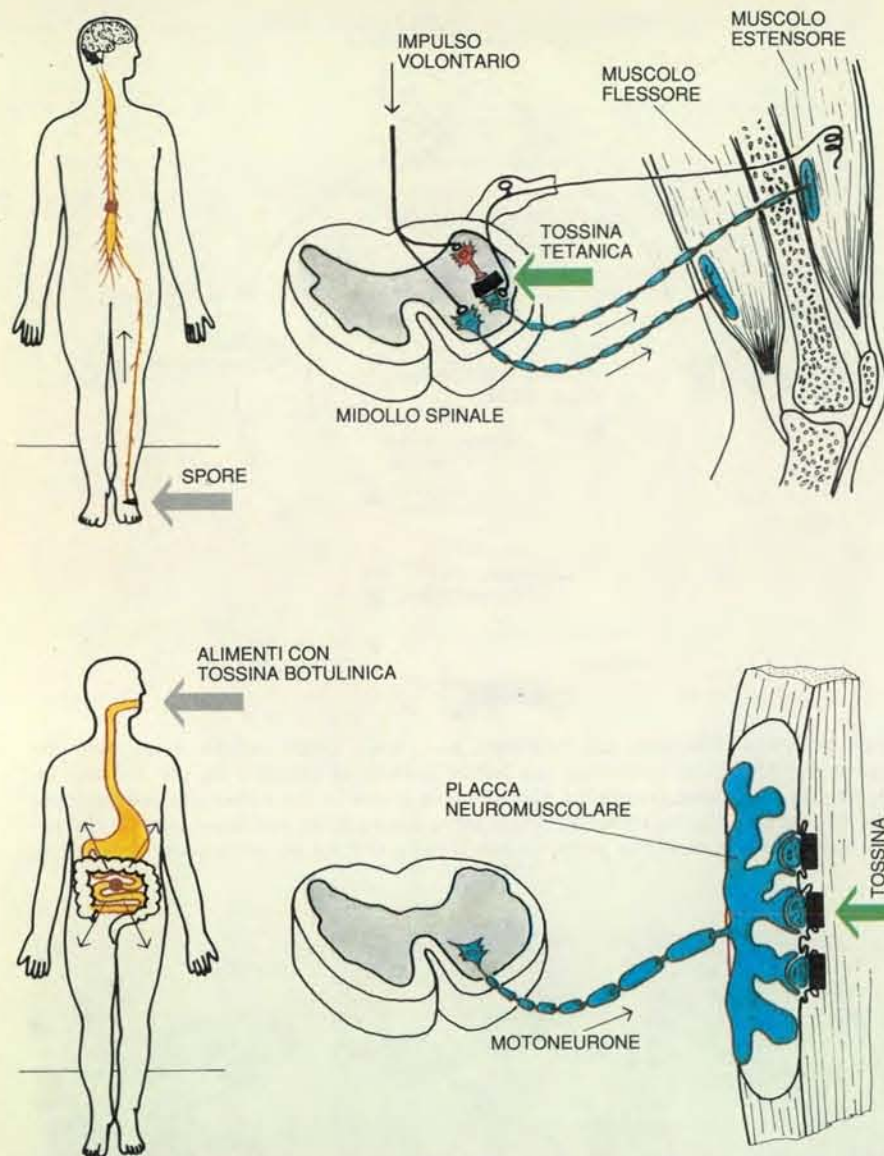
Una forma particolarmente subdola di botulismo è quella che colpisce i bambini nei primi mesi di vita e per questa sua elusività è stata scoperta solo nel 1976 da Jackson Pickett, Bruce Berg e collaboratori dell'Università della California a San Francisco e da Thaddeus Midura e Steve Arnon del California State Department of Health di Berkeley. Nel botulismo infantile, spore di clostridi tossinogenici, assunte con cibi contaminati, germinano nel tratto intestinale del neonato in cui non si è ancora sviluppata una flora batterica intestinale competitiva e producono la tossina. Il botulismo infantile è quindi un'infezione seguita da un'intossicazione, e non una semplice intossicazione.

Il botulismo, molto più del tetano, è una rilevantissima causa di morte negli animali, tanto da assumere a volte la forma di un'epidemia. Il fenomeno è più vistoso negli uccelli. Sulla carne in putrefazione crescono facilmente sia clostridi produttori di tossina sia larve, sviluppatesi da uova depositate da insetti, le quali si imbevono di tossina senza esserne danneggiate. Di larve sono ghiotte diverse specie di uccelli che, cibandosene, si intossicano e muoiono di botulismo. Si crea così un nuovo terreno anaerobio ideale per la crescita di altri clostridi che a loro volta produrranno altra tossina, generando un circolo vizioso mortale in cui perdono la vita decine,

centinaia di migliaia di animali per singolo episodio. Dato lo scarso rilievo epidemiologico del botulismo umano, non si ritiene necessaria una vaccinazione, di cui sentono invece la necessità gli allevatori per prevenire i focolai epidemici che si manifestano talvolta negli allevamenti con gravissimi danni economici.

La scoperta che tetano e botulismo sono causati da tossine proteiche ha dato il via a un'intensa attività di ricerca che ha portato in primo luogo al chiarimento del loro meccanismo d'azione a livello dell'organismo, e, successivamente, ai livelli cellulare e molecolare. Già nel 1892 Alessandro Bruschettini, allievo di Tizzoni a Bologna, stabilì che la tossina tetanica raggiunge il sistema nervoso centrale mediante un movimento retrogrado lungo l'assone del motoneurone. Nel 1949 Burgen, Dickens e Zatman della Middlesex Hospital Medical School di Londra dimostrarono che le neurotossine botuliniche bloccano la liberazione di acetilcolina a livello della giunzione neuromuscolare. In seguito a questa scoperta, e grazie al lavoro di numerosi laboratori, si stabilì che anche la tossina tetanica si lega alla giunzione neuromuscolare. Al contrario delle tossine botuliniche, essa non agisce a livello della giunzione, ma risale il motoneurone e viene liberata nello spazio intersinaptico fra motoneurone e interneurone inibitorio, penetra in queste seconde cellule e ne blocca il rilascio di neurotrasmettitori.

Si spiegano così tutti i sintomi clinici delle due malattie. Nel botulismo si ha una paralisi flaccida, perché la tossina blocca la trasmissione dell'impulso nervoso a livello della giunzione neuromuscolare; nel tetano si ha una paralisi spastica, perché viene inibito il circuito nervoso spinale che assicura il corretto bilanciamento della contrazione dei muscoli contrapposti, che è essenziale per un movimento armonico e controllato delle articolazioni.

La scoperta del blocco della giunzione neuromuscolare da parte delle tossine botuliniche ne ha promosso un uso crescente come farmaci nel trattamento di patologie caratterizzate da iperattività della giunzione neuromuscolare. In un lavoro pionieristico pubblicato nel 1973 Alan Scott e collaboratori, dello Smith-Kettlewell Eye Research Institute di San Francisco, dimostrarono che si potevano selettivamente indebolire i muscoli oculari di scimmie inoculando localmente quantità minutissime di tossina botulinica. Questo intervento è stato esteso a varie patologie umane, è reversibile e ha una efficacia limitata ad alcuni mesi. Infatti, l'inoculazione di tossina botulinica equivale funzionalmente a una denervazione localizzata e, come questa, induce formazione di nuove giunzioni neuromuscolari adiacenti. L'inoculazione deve essere quindi ripetuta, ma una complicanza non infrequente

Negli uccelli la diffusione del botulismo può essere molto elevata. L'animale che muore per botulismo costituisce un fertile terreno di crescita sia per il batterio, che produce la tossina botulinica, sia per larve di insetti insensibili alla neurotossina i cui tessuti ne vengono imbevuti. Delle larve sono ghiotti numerosi uccelli, che così si intossicano e muoiono perpetuando il ciclo, che ha un andamento epidemico.

La serie di microfotografie elettroniche ad alto ingrandimento riguarda tre diverse giunzioni neuromuscolari di rana a livello della «zona attiva». Le preparazioni sono state stimolate elettricamente per due ore a 2 hertz in presenza di perossidasi di rafano, utilizzata come tracciante extracellulare. Gli spazi intersinaptici contengono densi depositi di prodotti di reazione della perossidasi. Le immagini mostrano tre diversi gradi di associazione tra vescicole sinaptiche, marcate con perossidasi, e membrana presinaptica. In *a*, la membrana della vescicola è completamente fusa con la membrana presinaptica; in *b*, la continuità delle due membrane è mantenuta da uno stretto processo; in *c*, la vescicola sembra essere sul punto di perdere contatto con l'assolemma. (Le immagini sono tratte da Bruno Ceccarelli e altri, «The Journal of Cell Biology», 54, pp. 30-38, 1972, per cortesia di Flavia Valtorta del Dipartimento biotecnologie dell'Istituto di ricovero e cura San Raffaele di Milano.)

Sono schematizzati il meccanismo della genesi della paralisi spastica tipica del tetano (*in alto*) e quello della paralisi flaccida provocata dal botulismo. La tossina tetanica penetra a livello della giunzione neuromuscolare, ma, invece di agire a livello periferico, risale all'interno di vescicole l'assone del motoneurone. La tossina tetanica viene poi scaricata nello spazio intersinaptico a livello del midollo spinale e penetra negli interneuroni inibitori bloccandone il rilascio di neurotrasmettitori. Ciò causa l'inibizione del meccanismo di bilanciamento della contrazione dei muscoli contrapposti che è controllato da questi interneuroni, dando origine a una paralisi spastica. Le neurotossine botuliniche penetrano anch'esse nel motoneurone a livello della giunzione neuromuscolare e inibiscono in questa sede la liberazione di acetilcolina, bloccando così la trasmissione dell'impulso nervoso al muscolo.

delle inoculazioni ripetute è la produzione di anticorpi antitossina. Questo problema può comunque venire superato grazie all'impiego di sierotipi differenti. Il National Institute of Health di Washington ha dichiarato recentemente che il trattamento con tossina botulinica è il metodo terapeutico d'elezione per una serie di distonie come blefarospasmo, strabismo, torcicollo, disfonia. Le tossine botuliniche trovano inoltre un uso crescente nel trattamento di patologie dei grandi muscoli dorsali e degli arti. È sorprendente vedere come, grazie alla ricerca scientifica, in pochissimo tempo, si sia passati dalla scoperta delle più potenti tossine note al loro impiego terapeutico.

Per quanto i sintomi clinici causati dalle neurotossine del tetano e del botulismo siano opposti, esse agiscono nello stesso modo, vale a dire bloccando la liberazione di neurotrasmettitori da parte dei neuroni. Quest'unificazione del meccanismo d'azione a livello cellulare ha stimolato la ricerca verso due obiettivi principali: scoprire il meccanismo d'azione delle neurotossine a livello molecolare e saperne di più sulla neuroesocitosi, il processo mediante il quale il neurotrasmettitore, o il neuro-ormone, contenuto all'interno di vescicole sinaptiche viene liberato nello spazio intersinaptico oppure nella giunzione neuromuscolare.

Si è così trovato che queste tossine vengono prodotte da varie specie di clostridi in forma di singola catena polipeptidica di 150 000 dalton, priva d'attività. Successivamente, proteasi batteriche o tissutali tagliano in un singolo punto la tossina trasformandola nella forma bicatenaria: una catena pesante (*H*, 100 000 dalton) e una catena leggera (*L*, 50 000 dalton) tenute assieme da un singolo ponte disolfuro intercatena e da forze non covalenti.

I dati disponibili e il confronto con il processo di intossicazione cellulare della tossina difterica, studiato di recente in dettaglio nel nostro laboratorio da Emanuele Papini, suggeriscono che l'intossicazione neuronale proceda attraverso un meccanismo a quattro stadi: *a*) legame a cellule neuronali; *b*) penetrazione cellulare all'interno di vescicole; *c*) traslocazione attraverso la membrana vescicolare, riduzione del ponte disolfuro e liberazione della catena *L* nel citoplasma; *d*) blocco della neuroesocitosi. La catena *H* è principalmente responsabile del legame specifico con le cellule neuronali, anche se non sono ancora state identificate le molecole della membrana plasmatica coinvolte nel processo. Questo è un aspetto particolarmente importante perché si pensa sia alla base della variabilità della risposta dei pazienti all'inoculazione con tossina botulinica e della differente tossicità di queste neurotossine nelle varie specie animali.

Di recente abbiamo dimostrato che cellule di feocromocitoma di ratto fatte differenziare in senso neuronale con NGF (fattore di crescita neuronale) esprimono sulla superficie cellulare una proteina di circa 20 000 dalton che interagisce specificamente con la catena *H* della tossina tetanica. Altri dati indicano che, oltre a componenti proteici, anche lipidi acidici come i gangliosidi sono coinvolti nel legame cellulare di quantità di tossina estremamente piccole, ma sufficienti a provocare la malattia.

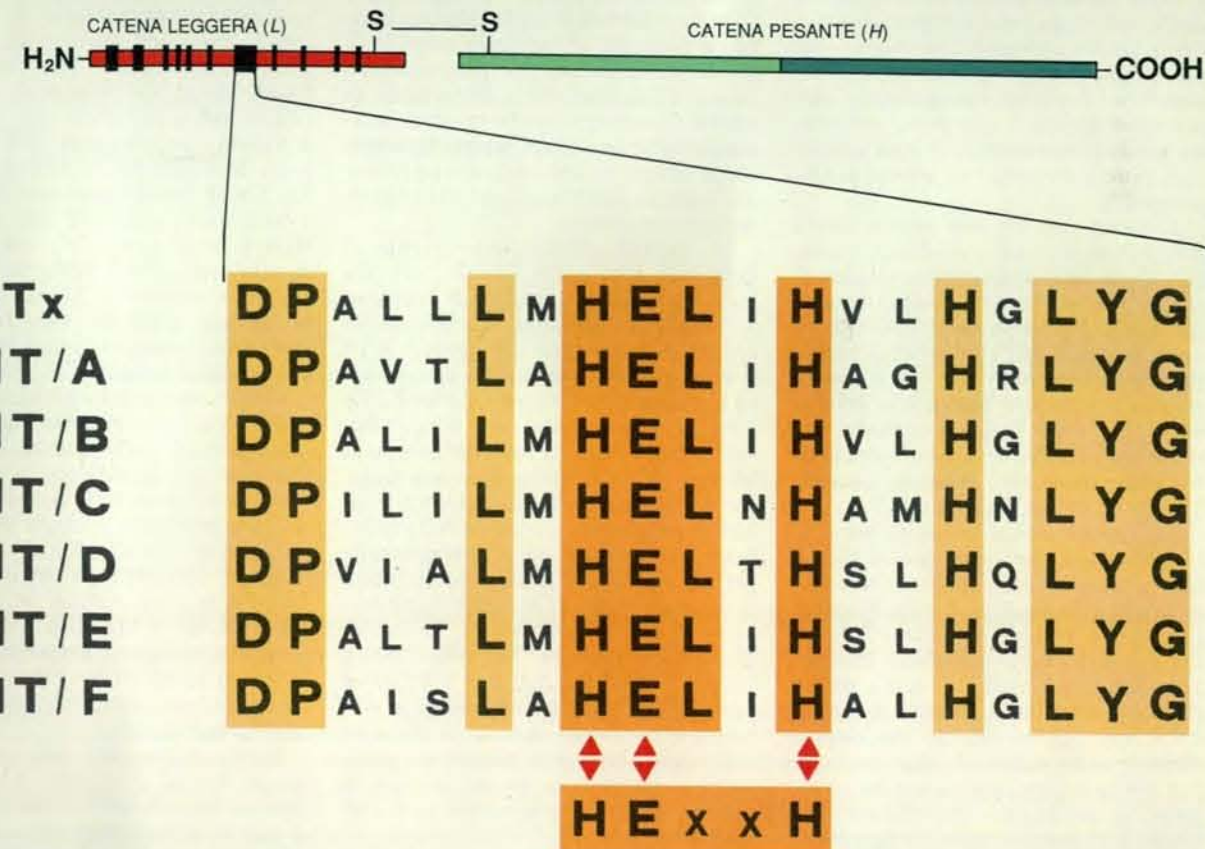
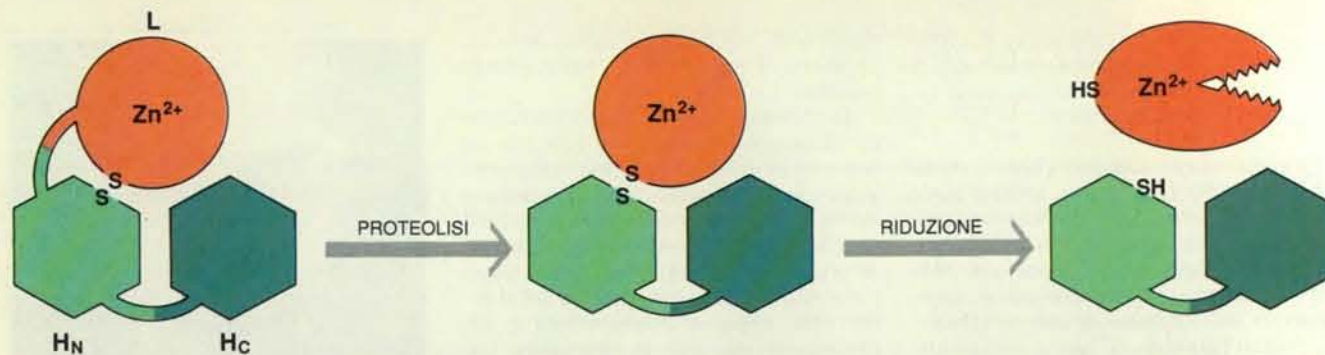
L'interazione con la membrana plasmatica è seguita dall'ingresso della tossina nella cellula tramite un processo d'internalizzazione in vescicole di natura ancora non nota. Elaine Neale del National Institute of Health di Bethesda ha recentemente fornito prove del fatto che la tossina penetra nel citoplasma a partire da vescicole, il cui interno sia reso acido da una pompa protonica ATPasica. Si pensa che la catena *H* cambi conformazione all'abbassarsi del *p*H, come è stato suggerito da Patrice Boquet dell'Istituto Pasteur di Parigi, e che assista la catena *L* nella sua traslocazione nel citoplasma. Ulteriori conoscenze su questo aspetto sono necessarie per lo sfruttamento di queste tossine come vettori di farmaci o altri agenti biologici all'interno delle cellule nervose.

Recentemente abbiamo chiarito come le catene *L* di queste tossine blocchino il rilascio di neurotrasmettitori. Abbiamo infatti dimostrato che le catene leggere di queste neurotossine sono delle zincoproteasi specifiche per le componenti dell'apparato di neuroesocitosi.

Siamo partiti da un'analisi comparativa delle sequenze amminoacidiche delle neurotossine prodotte da clostridi ottenute recentemente per clonaggio e sequenziamento dei relativi geni, grazie al lavoro compiuto da vari laboratori in tutto il mondo, tra cui quelli di Heiner Niemann all'Università di Tübingen, di Michel Popoff all'Istituto Pasteur di Parigi, di Nigel Minton nel Laboratorio di sanità pubblica di Salisbury e di Neil Fairweather dei Laboratori Wellcome. Contrariamente a quanto ci si aspettava, l'omologia fra le varie tossine è limitata solo ad alcuni segmenti non consecutivi e localizzati prevalentemente nella catena *L*, quella responsabile del blocco della neuroesocitosi. Fra questi, il segmento più conservato si trova al centro della catena *L* e contiene la breve sequenza istidina-glutammico-xaa-xaa-istidina, tipica delle zincoendopeptidasi, dove essa media il legame dello zinco con la proteina.

Le endopeptidasi zinco-dipendenti costituiscono una grande famiglia di proteasi di enorme interesse: zincoproteasi specifiche controllano la pressione arteriosa (convertono angiotensina I ed endotelina I nei corrispondenti peptidi ad azione vasocostrittrice), sono coinvolte nella degradazione delle encefaline, nella risposta infiammatoria e nei conseguenti processi di riparazione tissutale (proteasi di leucociti) e nei processi d'invasione e metastasi dei tumori maligni. Proprio a causa dei loro importanti e molteplici ruoli, di queste proteasi si conosce molto: dalla struttura tridimensionale di quattro di esse, al meccanismo di scissione del legame peptidico, in cui ha un ruolo fondamentale lo zinco, a centinaia di inibitori.

Successivamente a queste osservazioni iniziali, abbiamo misurato il contenuto di metalli pesanti della tossina tetanica e di cinque sierotipi di tossina botulinica. Abbiamo così trovato che tutte contengono un atomo di zinco per molecola di tossina legato alla catena *L* attraverso le due istidine del segmento centrale conservato, proprio come si prevedeva in base all'omologia con le zincoproteasi.

Una paziente affetta da blefarospasmo come appariva prima (*in alto*) e dopo il trattamento con tossina botulinica effettuato da Emilio Campos della Clinica oculistica dell'Università di Modena.

Nello schema della struttura delle neurotossine del tetano e del botulismo e del segmento omologo (amminoacidi indicati con il codice a lettera singola) è evidenziato (*in arancione*) il motivo di legame dello zinco delle zincoendopeptidasi costituito da istidina-glutammico-xaa-xaa-istidina (HExxH, nel codice a lettera singola degli amminoacidi riportato nell'illustrazione).

Nel processo di attivazione delle neurotossine prodotte da clostridi, queste sono sintetizzate in forma di singola catena (*a sinistra*), poi convertite mediante proteolisi nella forma bicatenaria (*al centro*) e attivate mediante riduzione, all'interno del neurone, del singolo ponte disolfuro intercatena (*a destra*). Queste proteine sono costituite da tre domini strutturalmente e funzionalmente distinti: la catena (*in rosso*) *L* è responsabile del blocco della neuroesocitosi; la porzione carbossiterminale della catena pesante H_C è principalmente responsabile del legame specifico con i neuroni, mentre la porzione amminoterminale della catena pesante H_N è coinvolta nella traslocazione di *L*, attraverso la membrana cellulare, nel citoplasma.

Questi primi risultati ci hanno spinto ad approfondire il ruolo dello zinco della tossina nel processo di intossicazione dei neuroni. In una serie di esperimenti di microiniezione di tossine, variamente modificate, in coppie di neuroni del ganglio buccale di *Aplysia californica*, un mollusco marino, svolti in collaborazione con Bernard Poulain del Laboratorio di neurobiologia cellulare e molecolare del CNRS a Gif-sur-Yvette, si è trovato che il blocco della neuroesocitosi da parte della tossina tetanica è strettamente dipendente dalla presenza dell'atomo di zinco. Inoltre abbiamo dimostrato come gli inibitori delle zincoendopeptidasi blocchino l'attività intracellulare della tossina fornendo una prova indiretta, ma molto convincente, che essa intossica la cellula attraverso un'attività di tipo proteasico.

Abbiamo fornito una prova diretta dell'attività proteasica mediante l'identificazione del bersaglio intracellulare di queste neurotossine. In collaborazione con Fabio Benfenati dell'Istituto di fisiologia umana dell'Università di Modena, abbiamo infatti trovato che la tossina tetanica e le tossine botuliniche dei sierotipi B, D ed F tagliano in modo specifico una proteina della membrana delle vescicole sinaptiche, chiamata sinaptobrevina (o anche VAMP).

La sinaptobrevina è una proteina di 13 000 dalton, che, come spesso accade dopo l'avvento delle moderne tecniche di genetica molecolare, è stata scoperta, da William Trimble e Richard Scheller dell'Howard Hughes Medical Institute della Stanford University, non a partire dal materiale proteico, bensì a partire da RNA messaggero estratto dall'organo elettrico di torpedine. Il gene che codifica per la sinaptobrevina è stato in seguito clonato da ratto, da moscerino, da bovino e da uomo e ne è stata dimostrata l'espressione nel tessuto nervoso e neuroendocrino da parte di Francesca Navone e Pietro De Camilli, del Dipartimento di farmacologia dell'Università di Milano, e da Reinhard Jahn, allora al Max-Planck-Institut di Monaco.

L'identificazione della sinaptobrevina come comune sito d'azione intraneuronale della neurotossina tetanica e delle neurotossine botuliniche di sierotipo B, D ed F unifica a livello molecolare il meccanismo d'azione di queste tossine, dopo l'unificazione a livello cellulare effettuata nei primi anni cinquanta. Fra l'altro, ciò ha notevoli implicazioni per gli studi volti a trovare inibitori specifici di queste tossine da impiegarsi come farmaci antitetano e antibotulismo. Inoltre, il fatto che il bersaglio comune di queste neurotossine, che provocano un blocco persistente della liberazione di neurotrasmettitori, sia una proteina localizzata sulle vescicole sinaptiche costituisce una prova forse definitiva a favore dell'ipotesi vescicolare del rilascio di neurotrasmettitori.

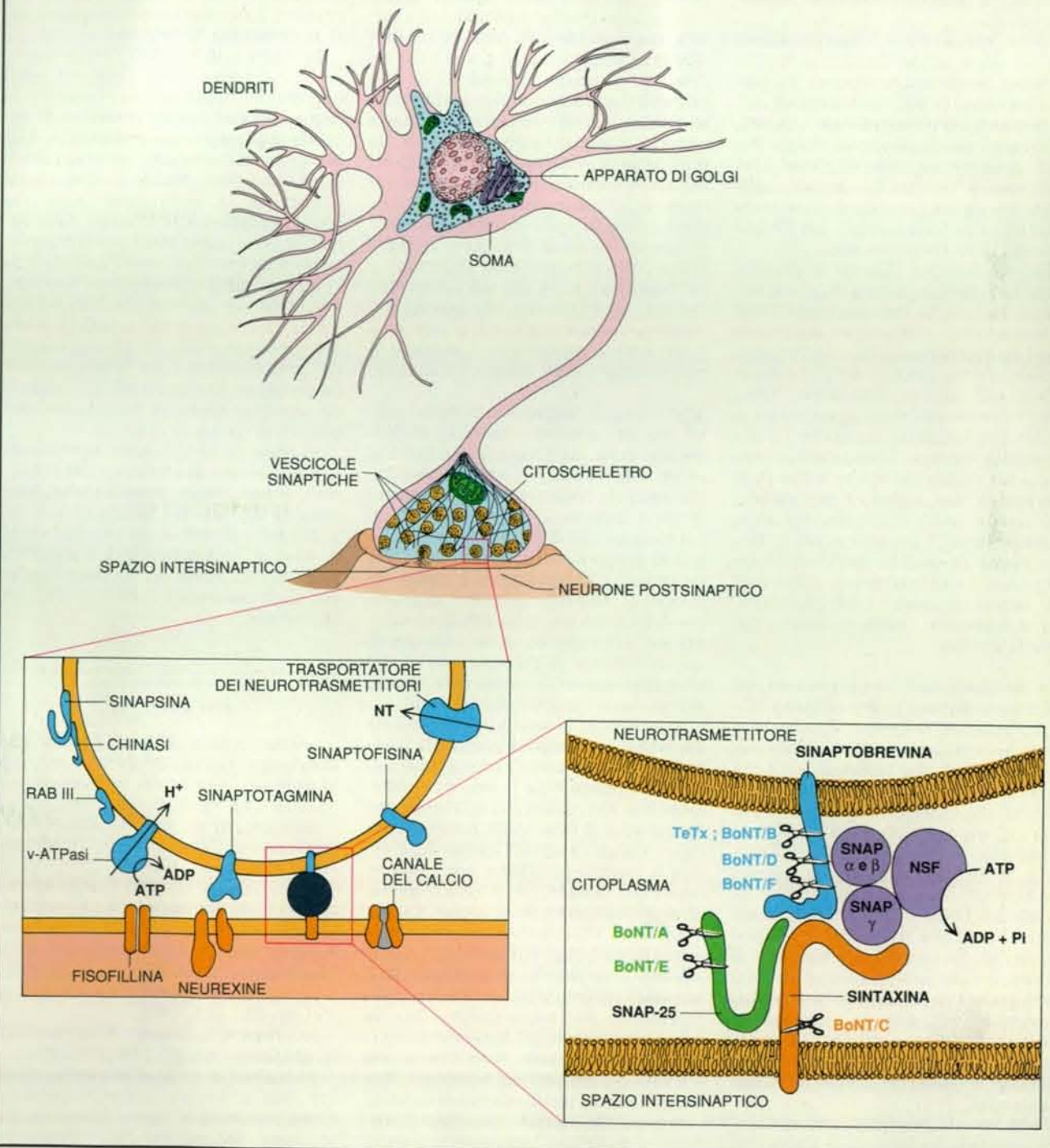
Di recente, contemporaneamente ai gruppi di Jahn e De Camilli, ora alla Yale University, e a quello di Niemann a Tübingen, abbiamo identificato in SNAP-25 il bersaglio dei sierotipi A ed E e in sintaxina quello della neurotossina botulinica di sierotipo C. SNAP-25 è stata scoperta nel 1989 nello Scripps Research Institute di La Jolla dal gruppo di Michael Wilson, sempre mediante tecniche di biologia molecolare. Essa è localizzata sulla faccia citosolica della membrana presinaptica ed è indispensabile per la crescita dell'assone, come hanno dimostrato Stefano Catsicas e collaboratori del Glaxo Institute for Molecular Biology di Ginevra. Nel 1992, Mark Bennet e Richard Scheller a Stanford hanno identificato e sequenziato la sintaxina e ne hanno rilevato la presenza nelle zone attive della membrana presinaptica, quelle zone in cui avviene la fusione della vescicola carica di neurotrasmettitore.

Sempre di recente, per merito di James Rothman, Thomas Sollner e collaboratori del Sloan-Kettering Institute di New York, la linea di ricerca sulla patologia molecolare del tetano e del botulismo si è fusa con una indagine genetica del processo di secrezione, svolta in lievito dai gruppi di Randy Schekman dell'Università della California a Berkeley e di Peter Novick della Yale University, e con una linea di ricerca sulla biochimica del processo di fusione di membrane cellulari, perseguita in molti laboratori diversi. Rothman e collaboratori hanno dimostrato che il processo mediante il quale vescicole cellulari riconoscono il loro compartimento di destinazione e ne fondono le rispettive membrane viene mediato da un complesso multiproteico. Tale complesso è formato da alcune proteine solubili identificate nel corso degli anni ottanta da Rothman e collaboratori: NSF, una proteina dotata di attività ATPasica e i fattori SNAP che esistono nelle forme alfa e beta, molto simili tra loro, e gamma. Inoltre fanno parte del complesso sinaptobrevina, che è localizzata sulle vescicole, e sintaxina e SNAP-25, che sono invece ambedue localizzate sulla membrana bersaglio con cui la vescicola è destinata a fondersi.

Il fatto che tre delle proteine proposte come parte del complesso multiproteico siano bersagli delle neurotossine indica che esse ne costituiscono il nucleo essenziale ed è sinora la prova sperimentale più importante a favore del modello di Rothman. Ma altri dati si stanno rapidamente accumulando. In questa luce non c'è da meravigliarsi se, assieme a Luisa Gorza e Maurizio Vitadello del nostro Dipartimento, abbiamo adesso localizzato questa proteina in molti tessuti, e non solo nel tessuto nervoso come era noto in precedenza.

La sinaptobrevina viene tagliata da queste forbici molecolari a zinco, che sono le neurotossine, in un solo punto: il singolo legame peptidico fra glutammina e fenilalanina nel caso delle tossine tetanica e botulinica B e fra glutammina e lisina se si tratta del sierotipo F, mentre il sierotipo D idrolizza un singolo lega-

Le proteine bersaglio delle neurotossine prodotte dai clostridi

La sinaptobrevina (o VAMP) svolge un ruolo fondamentale nel legame delle vescicole sinaptiche con le «zone attive» della membrana presinaptica mediante il riconoscimento specifico di SNAP-25 e di sintaxina. In seguito a tale riconoscimento, si assembla un complesso cui partecipa una serie di proteine citosoliche, tra cui i fattori di Rothman: una proteina ad attività ATPasica sensibile alla N-etil-maleimmide (NSF) e due proteine solubili e accessorie di NSF: SNAP alfa o beta e gamma. In modo del tutto ignoto, segue la fusione della membrana della vescicola con la membrana plasmatica, l'evento che porterà al rilascio del neurotrasmettitore. Questo processo viene fatto scattare, nello spazio di un centinaio di microsecondi, dall'entrata di ioni calcio attraverso i canali del calcio, che si trovano nelle vicinanze e che vengono aperti dalla depolarizzazione della membrana plasmatica. Come si vede nell'ingrandimento a destra, le neurotossine del tetano e del botulismo bloccano la neuroesocitosi mediante un singolo specifico taglio di tre componenti essenziali a questo processo. La tossina tetanica e le tossine botuliniche dei sierotipi B, D ed F tagliano la sinaptobrevina, legata alle vescicole che contengono il neurotrasmettitore. Al contrario due proteine della membrana presinaptica vengono tagliate dai sierotipi A ed E, che riconoscono specificamente la SNAP-25, e dalla tossina botulinica di sierotipo C, che agisce sulla sintaxina.

me peptidico lisina-leucina di sinaptobrevina. La neurotossina botulinica di sierotipo C taglia la sintaxina vicino alla sua àncora di membrana. Anche i sierotipi A ed E tagliano SNAP-25 in un singolo punto, che è diverso per le due neurotossine. Nel caso di sinaptobrevina e di sintaxina, la tossina stacca dalle corrispondenti membrane una parte rilevante del loro dominio citosolico con il risultato che la vescicola non può più fondere con la membrana presinaptica e scaricare il proprio contenuto di neurotrasmettitori all'esterno. L'effetto è il blocco della neuroesocitosi con le conseguenze drammatiche sopra descritte, proprio a causa del ruolo essenziale del processo di neuroesocitosi nella fisiologia animale.

Meno immediata è l'interpretazione dell'azione su SNAP-25, da cui la neurotossina botulinica A rimuove un piccolo segmento di nove amminoacidi dalla porzione carbossiterminale. Quindi, l'eliminazione di meno del cinque per cento della massa totale di SNAP-25 è sufficiente a bloccare la funzione, indicando che questa porzione della molecola ha un ruolo fondamentale nel funzionamento della neuroesocitosi.

Queste scoperte valgono a spiegare anche la tremenda potenza di queste tossine. Infatti, se la loro neurospecificità ha una funzione primaria nel determinare la loro estrema tossicità, si deve anche considerare che una sola molecola di catena *L* può tagliare, una dopo l'altra, tutte le molecole di sinaptobrevina o SNAP-25 o sintaxina, presenti nella terminazione nervosa. È impossibile pensare a un veleno che agisca a una dose inferiore a una molecola per cellula. Ciò spiega anche la persistenza della malattia. Infatti è necessario che la tossina venga in qualche modo inattivata prima che la cellula possa ricostituire per sintesi *ex novo* il proprio patrimonio di proteine sinaptiche danneggiate dalla tossina.

La sinaptobrevina viene prodotta dal tessuto nervoso in due isoforme (1 e 2), le quali differiscono significativamente solo nella parte amminoterminale. Il ruolo delle due isoforme non è noto. È interessante notare come sia la neurotossina tetanica sia quella botulinica B taglino, nel ratto e nel pollo, solo la isoforma 2. Grazie al lavoro di Trimble e Scheller, di Thomas Südhof del Southwestern Medical Center dell'Università del Texas a Dallas e di Tomaso Patarnello e Luca Bargelloni del Dipartimento di biologia dell'Università di Padova, è stato dimostrato che nel ratto e nel pollo l'isoforma 1 differisce dall'isoforma 2 nel sito di proteolisi: la glutammina è stata sostituita da una valina. Al contrario, nell'uomo e nel topo ambedue le isoforme contengono una glutammina.

Sulla base di questi dati, abbiamo proposto che la bassa tossicità di queste tossine nel ratto e nel pollo e l'alta tossicità nel topo e nell'uomo possa essere, almeno in parte, dovuta proprio a questa singola mutazione occorrente nel sito di taglio delle tossine e non solamente alla presenza o meno a livello della giunzione neuromuscolare di recettori ad alta affinità per le neurotossine.

Un aspetto rilevante dei nostri studi è che l'attività di queste forbici molecolari a zinco sia bloccata da peptidi che contengono la sequenza di taglio e da inibitori di zincoproteasi come il captopril, un farmaco anti-ipertensivo largamente usato, che inibisce specificamente la zincoproteasi deputata alla conversione dell'angiotensina I in angiotensina II, un potente vasocostrittore. Questi agenti sono efficaci *in vitro* o iniettati nella cellula, ma, dato che non sono permeabili alla membrana plasmatica, non sono efficaci se vengono aggiunti al liquido extracellulare. Sebbene di scarso interesse clinico, questi risultati sono tuttavia importanti perché fanno intravvedere la possibilità di sviluppare nuovi composti che siano inibitori delle neurotossine del tetano e del botulismo e al tempo stesso siano permeabili alle membrane cellulari. Questi composti potrebbero rivelarsi efficaci per il trattamento del tetano e del botulismo. Oltre che nel trattamento dei pazienti tetanizzati, tali farmaci potrebbero risultare molto utili ai neurologi e agli oculisti per risolvere i problemi di sovradosaggio delle tossine botuliniche.

Se i risultati sopradescritti hanno portato un contributo essenziale al chiarimento della patogenesi molecolare del tetano e del botulismo, rimangono tuttavia molti altri punti oscuri. Quali sono i recettori delle varie neurotossine? Qual è il meccanismo di penetrazione cellulare e di trasporto retroassonico della tossina tetanica? Rispondendo a queste domande, si scopriranno nuovi meccanismi della fisiologia della cellula nervosa. Ma un altro aspetto della biologia di questo sistema si presenta ancora più sfuggente: quello del ruolo delle tossine nell'ecologia batterica dei clostridi.

In generale, le tossine sono prodotte dai batteri allo scopo di alterare la fisiologia dell'ospite così da creare condizioni tali da promuovere la loro moltiplicazione e la loro diffusione nell'ambiente. Se questo è il ruolo delle tossine batteriche, l'alterazione dell'ospite deve essere la minima possibile e non causarne comunque la morte: un evento contrario al raggiungimento dello scopo. Questa condizione viene soddisfatta in tutte le infezioni da batteri produttori di tossine: soltanto una piccola percentuale degli ammalati muore infatti di difterite, colera, pertosse, dissenterie da *Shigella* o da ceppi tossinogenici di *Escherichia coli*.

Come già abbiamo visto, questo non è il caso del tetano e del botulismo. Anche nella quasi totale assenza di dati sull'ecologia dei clostridi, comunque di difficile acquisizione, possiamo tentare di fornire una spiegazione, partendo naturalmente dal punto di vista del batterio anaerobio e non dal punto di vista delle potenziali vittime dell'infezione. In generale gli animali non contengono zone di anaerobiosi tali da favorire la crescita di clostridi, a parte porzioni ridotte dell'intestino di alcune specie. Così, la morte dell'animale, e non solo la necrosi di alcune aree tissutali ristrette, come avviene nella difterite, fornisce una condizione sicuramente vantaggiosa per un organismo anaerobio perché ne permette la crescita sulla carne in putrefazione: in parole povere, l'involontario ospite animale viene trasformato dai clostridi produttori di tossina in un terreno anaerobio ottimale per la loro moltiplicazione. Più grande è l'animale ucciso, maggiore sarà la possibilità di crescita.

La «scelta» di uccidere l'ospite sarebbe quindi forzata dalle necessità vitali opposte dell'ospite e del clostridio: la prima necessariamente aerobia e la seconda obbligatoriamente anaerobia, dato che i clostridi non sono in grado di ridurre l'ossigeno. Inoltre si deve considerare che la trasmissione attraverso spore dei clostridi li affranca dalla necessità di un ospite vitale per la trasmissione della specie (si veda l'articolo *L'evoluzione della virulenza* di Paul W. Ewald in «Le Scienze» n. 298, giugno 1993). Di conseguenza, a qualche punto del processo evolutivo, l'acquisizione del gene codificante per le neurotossine ha permesso la transizione dei clostridi da saprofagi passivi a batteri attivi nel procurarsi nuove risorse.

In quest'ottica diventano significativi i primi sintomi del tetano e del botulismo. Infatti, anche le alterazioni lievi come la diplopia nel botulismo o la rigidità del collo nel tetano, conducono, al di fuori del laboratorio o dell'ospedale, a una rapida morte per diminuita capacità di procacciarsi il cibo o di sfuggire ai predatori.

BIBLIOGRAFIA

MONTECUCCO C., *How do Tetanus and Botulinum Toxins bind to Neuronal Membranes?* in «Trends in Biochemical Science», 11, pp. 314-318, 1986.

SIMPSON L. L. (a cura), *Botulinum Neurotoxins and Tetanus Toxin*, Academic Press, San Diego, 1989.

SCHIAVO G., BENFENATI F., POULAIN B., ROSSETTO O., POLVERINO DE LAURETO P., DASGUPTA B. R. e MONTECUCCO C., *Tetanus and Botulinum B Neurotoxins block Neurotransmitter Release by Proteolytic Cleavage of Synaptobrevin* in «Nature», 359, pp. 832-835, 1992.

HUTTNER W. B., *Snappy Exocytotoxins* in «Nature», 365, pp. 104-105, 1993.

PATARNELLO T., BARGELLONI L., ROSSETTO O., SCHIAVO G. e MONTECUCCO C., *Neurotransmission and Secretion* in «Nature», 364, pp. 581-582, 1993.

Fluidi elettroreologici

Questi materiali liquidi, dotati della straordinaria proprietà di passare istantaneamente allo stato solido per esposizione a campi elettrici, potrebbero consentire la costruzione di macchine più veloci e versatili

di Thomas C. Halsey e James E. Martin

Il T-1000 - robot quasi indistruttibile fra i protagonisti del film *Terminator 2: Il giorno del giudizio* - può trasformarsi con la massima facilità da liquido a solido. Quando viene perforato da un proiettile, il rivestimento di metallo liquido del robot scorre fino a riparare il foro; quando viene fatto a pezzi il T-1000 fonde e poi si ricostituisce nella sua forma originale. I fluidi elettroreologici potrebbero avvicinarsi alla strabiliante adattabilità immaginata dai produttori del film.

Un fluido elettroreologico è una sostanza la cui viscosità cambia in presenza di campi elettrici. A seconda della forza del campo a cui è soggetto, esso può scorrere liberamente come acqua, fluire più lentamente come miele o assumere la solidità della gelatina, e il passaggio da uno stato all'altro può avvenire in pochi millisecondi. I fluidi elettroreologici sono facili da preparare: consistono semplicemente di particelle microscopiche sospese in un liquido isolante. Attualmente, però, non sono ancora abbastanza perfezionati per la maggior parte delle applicazioni commerciali; fra i problemi che presentano vi sono la debolezza strutturale allo stato solido, le proprietà abrasive allo stato liquido e l'instabilità chimica, soprattutto alle alte temperature.

Di recente si è riusciti a preparare fluidi elettroreologici con ingredienti meno abrasivi e una maggiore stabilità chimica, ma la debolezza strutturale rimane un serio inconveniente. Per poter progettare fluidi elettroreologici di impiego pratico, occorrerà una solida conoscenza delle interazioni elettriche fra particelle in sospensione che possono dare origine a solidi più resistenti.

I fluidi elettroreologici affascinano gli scienziati non solo a causa delle loro utili proprietà fisiche, ma anche per la loro complessità. Le particelle che li costituiscono si organizzano in una grande varietà di intricate strutture: a un estremo, quando il fluido scorre liberamente, le particelle si muovono in maniera indipendente; all'altro estremo, quando il fluido solidifica, esse si uniscono a formare sottili catene o colonne più spesse che sono visibili a occhio nudo.

Nello studio dei fluidi elettroreologici analizziamo dunque il comportamento della materia in una situazione intermedia fra i normali stati solido e liquido, collocandoci in una posizione abbastanza distante dalla tradizionale fisica dello stato liquido e dello stato solido. Recenti ricerche hanno approfondito le nostre conoscenze sulle interazioni fra particelle in questi stati peculiari; resta da vedere se queste scoperte consentiranno di ottenere fluidi elettroreologici con maggiore resistenza strutturale.

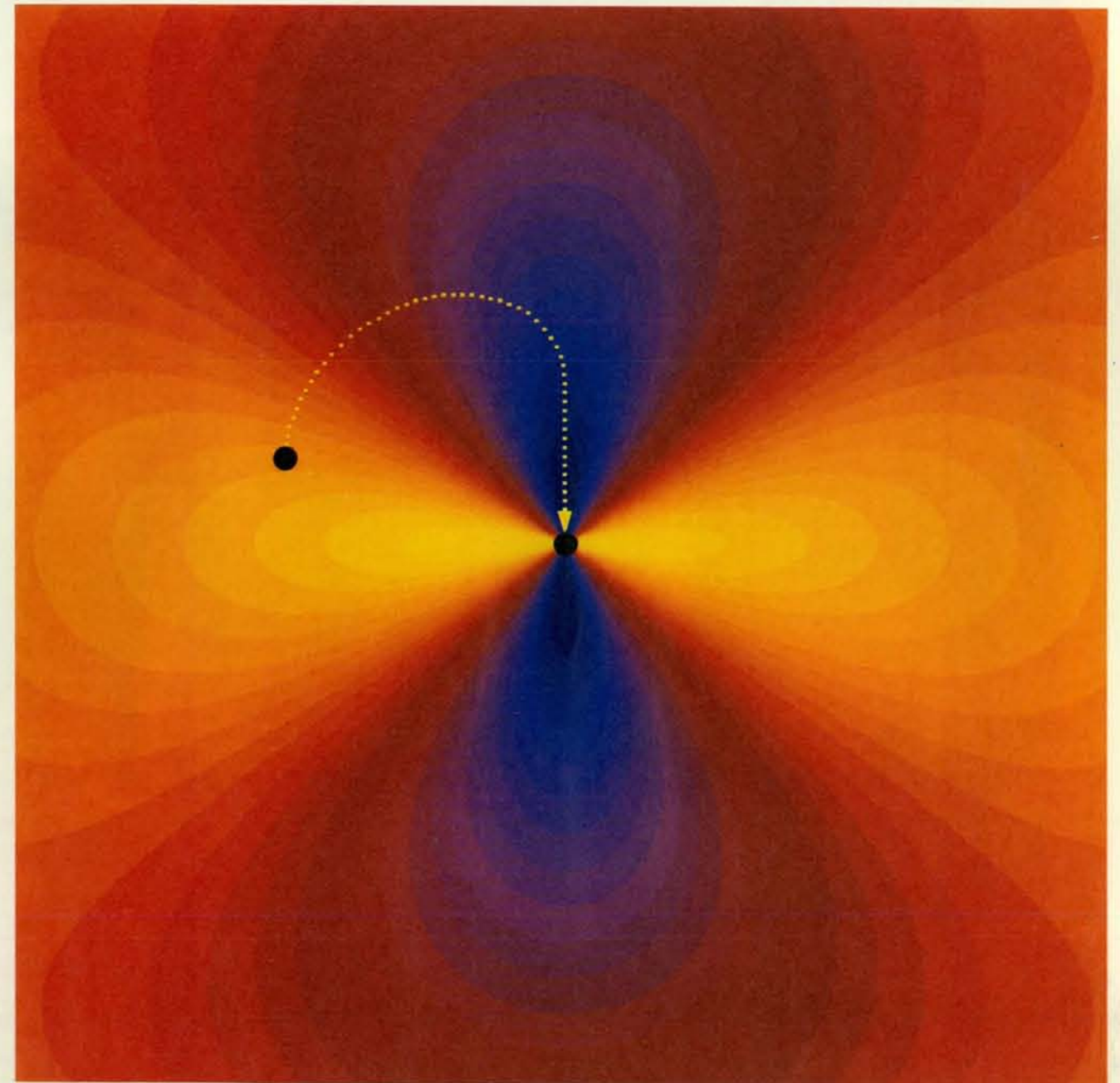
L'interesse per l'effetto elettroreologico nacque intorno al 1947, quando Willis M. Winslow, dopo circa otto anni di osservazioni, ottenne il primo brevetto per un fluido elettroreologico. Si scoprì ben presto che per ottenere un buon fluido elettroreologico bastava disperdere amido di mais in olio di mais. Questa preparazione «casalinga» presenta l'effetto elettroreologico quando è sottoposta a un intenso campo elettrico, pari a circa 1000 volt per millimetro. L'effetto può essere prodotto, per esempio, ponendo il fluido fra due lastre di metallo distanti un millimetro e caricando gli elettrodi fino a 1000 volt. Questa tensione può apparire molto alta, ma l'energia richiesta in realtà non è molta. Dato che il fluido è isolante, attraverso di esso non fluisce una corrente elettrica apprezzabile, e la carica delle lastre si accumula rapidamente fino a dare una tensione assai alta, allo stesso modo in cui migliaia di volt di elettricità statica possono accumularsi facilmente sulla pelle di una persona.

Il fluido di Winslow era una scoperta molto promettente, e il futuro dei fluidi elettroreologici appariva luminoso. Si poteva immaginare, per esempio, di costruire un'automobile dotata di frizione elettroreologica: il fluido della frizione avrebbe collegato il motore dell'automobile all'albero che trasmette il movimento alle ruote. Allo stato solido il fluido avrebbe costretto l'albero a ruotare, mentre allo stato liquido avrebbe permesso al motore di disimpegnarsi dall'albero e di ruotare liberamente, come quando è in folle. Una frizione siffatta avrebbe ben poche parti soggette a deterioramento o guasto e, con un tempo di risposta di qualche millisecondo, farebbe apparire lento un dispositivo puramente meccanico. Si sperava anche che i fluidi elettroreologici potessero costituire la base di sistemi idraulici privi di valvole, nei quali il fluido, solidificandosi, potesse impedire il flusso attraverso una tubazione sottile.

Le interazioni fra particelle cariche sono alla base dell'effetto elettroreologico. Una particella che ha un polo carico positivamente e uno carico negativamente può o respingere o attrarre una seconda particella polarizzata allo stesso modo (*qui sopra*). Quando le due particelle si trovano fianco a fianco, la repulsione fra poli uguali le fa allontanare, mentre quando sono allineate verticalmente l'attrazione fra poli diversi tende ad avvicinarle. Nell'illustrazione della pagina a fronte una particella polarizzata che venga a trovarsi nella zona di repulsione (*in arancione*) di un'altra particella si sposta fino a raggiungere la zona di attrazione (*in viola*). Se molte particelle polarizzate in sospensione in un fluido interagiscono in questo modo, si formano catene di particelle e il fluido manifesta l'effetto elettroreologico, ossia diventa più denso o perfino solido.

Queste idee sono interessanti ancora oggi, ma le possibilità davvero entusiasmanti riguardano dispositivi adattabili che sfruttino la viscosità continuamente variabile dei fluidi elettroreologici. Dato che la velocità di flusso di un fluido elettroreologico può essere controllata applicando un campo elettrico, dispositivi del tutto nuovi, quali ammortizzatori adattabili, hanno attirato l'interesse tanto dell'industria automobilistica quanto dei fautori dei treni a levitazione magnetica. (Anche se questi treni non corrono su rotaie, gli ammortizzatori adattabili potrebbero smorzare gli effetti di brusche partenze e frenate.)

Un comune ammortizzatore è costituito da un cilindro contenente un pistone che, quando si abbassa, pompa un olio denso e viscoso attraverso un piccolo orifizio. Il fluido assorbe la forza dell'impatto scorrendo lentamente, e torna poi più rapidamente nell'ammortizzatore durante la fase di richiamo, attraverso un orifizio più grande. Purtroppo la viscosità dell'olio dipende fortemente dalla temperatura. In una giornata fredda esso è più denso, e quindi l'ammortizzatore risulta un po' rigido. Inoltre, dato che dopo un certo numero di cicli l'olio si riscalda e diventa meno denso, la capacità del dispositivo di assorbire gli urti sparisce a poco a poco in un lungo tratto pieno di cunette, ossia proprio quando sarebbe più necessaria. Un modo per compensare queste difficoltà è quello di introdurre un sistema meccanico che regoli la dimensione dell'orifizio quando ciò sia necessario.

I fluidi elettroreologici consentirebbero di costruire ammortizzatori capaci di reagire in qualche millisecondo senza richiedere regolazioni meccaniche. Un dispositivo di questo genere può essere costruito riempiendo il cilindro con un fluido elettroreologico e aggiungendovi

elettrodi e un microprocessore. Quando il pistone si muove, il microprocessore modula rapidamente la tensione fra gli elettrodi per modificare la viscosità del fluido. Questo controllo veloce consentirebbe di aumentare la densità del fluido nel mezzo di un ciclo, in modo da smorzare l'effetto di una buca nella strada; il fluido potrebbe poi essere reso di nuovo meno denso per un rapido richiamo. Un unico ammortizzatore potrebbe così adattarsi a un gran numero di veicoli e di condizioni di funzionamento.

Tuttavia i fluidi elettroreologici attualmente esistenti non possono essere sottoposti a lavori troppo pesanti. Allo stato solido sono spesso cedevoli, e le tensioni elevate necessarie per farli solidificare, se applicate di frequente, possono riscaldare gli ingredienti liquidi volatili fino a provocarne l'evaporazione. Per risolvere questi problemi, stiamo studiando in che modo il campo elettrico causi la solidificazione. Idealmente, vorremmo disporre di fluidi che diventino solidi resistenti a tensioni più basse di quelle richieste oggi.

Si sa da qualche tempo che l'effetto elettroreologico prende origine dal fenomeno della polarizzabilità elettrica. In un campo elettrico - per esempio quello che si stabilisce fra una coppia di lastre di metallo cariche - i protoni contenuti in una particella, che hanno carica positiva, vengono attratti da uno degli elettrodi, mentre gli elettroni sono attratti dall'elettrodo opposto. Il risultato è che le cariche positive e negative migrano in una certa misura all'interno della particella, e in direzioni opposte. Una particella in un campo elettrico diventa così un dipolo elettrico, i cui poli negativo e positivo sono analoghi a quelli di un magnete. L'intensità di questo effetto varia a seconda della polarizzabilità elettrica della sostanza. Se questa è elevata, allora si ha una rilevante migrazione di cariche verso i poli opposti della particella; se la polarizzabilità è bassa, allora le particelle non reagiscono quasi al campo elettrico applicato.

Se il fluido contiene molte particelle polarizzate, le interazioni elettriche diventano complesse. La situazione più semplice è quella in cui due particelle sono immerse in un forte campo elettrico; poniamo che ciascuna particella abbia il polo positivo alla sua sommità e quello negativo nel punto più basso. La forza di attrazione o di repulsione fra queste due particelle dipende dal modo in cui esse si avvicinano l'una all'altra. Se la prima, stando direttamente sotto la seconda, si muove verso l'alto, allora il polo positivo dell'una è rivolto verso il polo negativo dell'altra e, dato che cariche opposte si attraggono, le due particelle risentono di una forza di attrazione.

Se la prima particella riesce in qualche modo ad affiancarsi alla seconda, allora sono i poli dello stesso segno a essere vicini, e fra le due particelle si esercita una forza di repulsione. Naturalmente, se le particelle si avvicinano l'una all'altra ad angolazioni intermedie, la forza di attrazione o di repulsione è minore che nei due casi estremi. In generale, se la congiungente le due particelle forma con la verticale un angolo inferiore a 55 gradi, esse si attraggono; in caso contrario si respingono.

Anche se la forza iniziale le respinge, le particelle tendono a muoversi in cerchio l'una intorno all'altra. Le loro traiettorie si incurvano, come quelle delle comete, in risposta alle forze di attrazione o repulsione. Nello stesso tempo, in qualunque direzione si stiano muovendo le particelle, i loro poli puntano sempre in su o in giù, verso gli elettrodi. Nel complesso, queste forze portano quasi inevitabilmente le particelle in una configurazione in cui i poli di carica opposta sono vicini; a questo punto le forze di attrazione prendono il sopravvento e le particelle aderiscono l'una all'altra. Particelle successive si allineano allo stesso modo, come le perle di una collana, fino a creare lunghe catene; questo comportamento è analogo a quello ben noto della limatura di ferro quando si dispone lungo le linee di forza di un magnete a barra. Le catene di particelle in un fluido elettroreologico si sviluppano rapidamente da un'estremità all'altra del contenitore; è la loro formazione a determinare la solidificazione del fluido.

Alcune particelle formano facilmente catene quando sono in sospensione in un tipo di fluido, ma non in un altro, e diversi ricercatori stanno ancora lavorando per comprendere i dettagli sfuggenti del processo. Una potenziale complicazione è il fatto che le particelle sono immerse in un liquido le cui molecole possono anch'esse polarizzarsi. Se la polarizzabilità del liquido è esattamente uguale a quella delle particelle, allora la sua polarizzazione maschera quella delle particelle e le catene non si formano. In generale, i migliori fluidi elettroreologici sono quelli in cui le particelle in sospensione si polarizzano più facilmente delle molecole del liquido.

Tuttavia le catene rappresentano solo uno stadio intermedio nell'evoluzione della struttura dei fluidi elettroreologici. Singole catene migrano lentamente verso catene parallele vicine e si combinano con queste a formare colonne più spesse, le quali si uniscono a costituire colonne ancora più spesse, e così via. Qual è la struttura solida finale? Di recente le ricerche teoriche e le osservazioni sperimentali hanno rivelato la risposta.

In primo luogo, un'analisi dei complessi campi elettrici associati a un gran numero di particelle offre indizi sulla struttura che ci si può attendere. Come abbiamo già visto, le particelle dipolari si respingono quando sono poste fianco a fianco, sicché la comparsa di colonne formate da catene parallele a prima vista appare misteriosa. Perché le colonne non si scompongono nelle catene che le costituiscono? Questo non accade perché ciò che attrae o respinge una particella rispetto a una catena vicina non è solo la forza esercitata dalle poche particelle limitrofe, ma anche quella generata da molte particelle lontane della catena. Le colonne stanno unite perché gli effetti della forza dipolare sono avvertiti fino a grande distanza, e le particelle reagiscono alla somma delle forze dovute a tutti gli altri dipoli del fluido.

Per comprendere perché le colonne non si scompongano, si consideri una particella «di prova» posta vicino a una catena. Essa viene respinta dalla particella più vicina della catena, come abbiamo visto nel caso di particelle messe fianco a fianco; viene però attratta dalle particelle situate molto al di sopra o al di sotto di essa, in quanto la congiungente fra queste e la particella di prova forma con la verticale un angolo inferiore a 55 gradi. Per essere più precisi, la particella di prova è attratta in una posizione che si trova leggermente al di sopra di una particella della catena e leggermente al di sotto di un'altra. Se molte particelle di prova sono poste vicino alla catena, esse risentono dell'attrazione di quest'ultima oltre che di quella reciproca e possono formare una seconda catena, poi una terza e così via. In questo modo le forze di dipolo che si esercitano fra particelle favoriscono la formazione di colonne in cui le particelle di catene adiacenti sono scalate l'una rispetto all'altra.

Tian-jie Chen, Robert N. Zitter e Rongjia Tao della Southern Illinois University hanno confermato di recente l'esistenza di questa struttura. Per il loro esperimento hanno utilizzato un fluido elettroreologico costituito da sferette di vetro di diametro relativamente grande - 40 micrometri - in sospensione in olio siliconico. Immediatamente dopo l'applicazione del campo elettrico si sono formate le catene, ed entro pochi minuti queste si sono aggregate in colonne di sezione grossolanamente circolare e di circa sei millimetri di diametro.

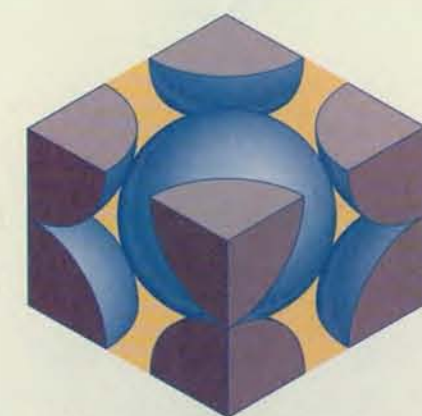
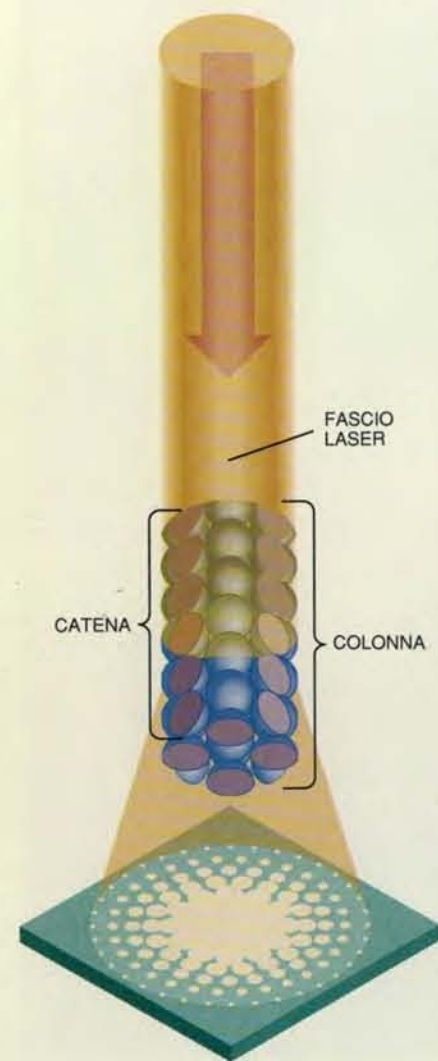
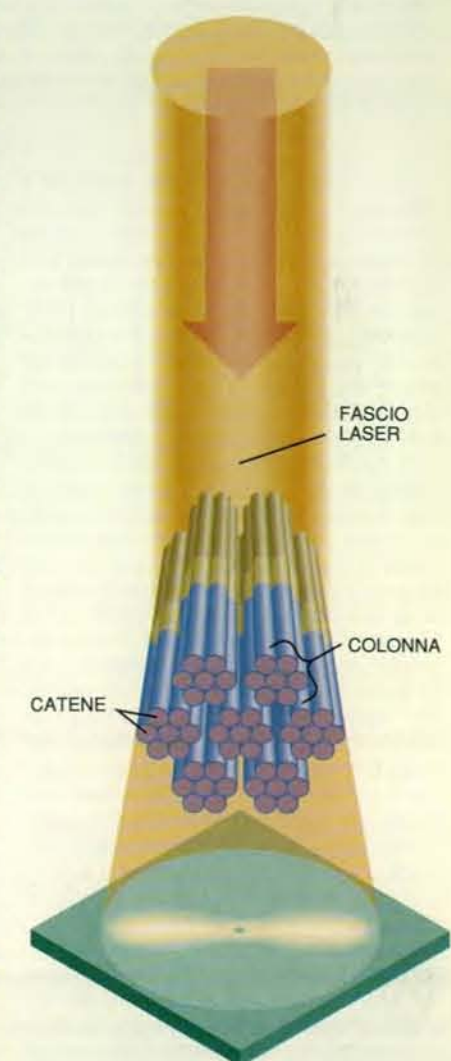
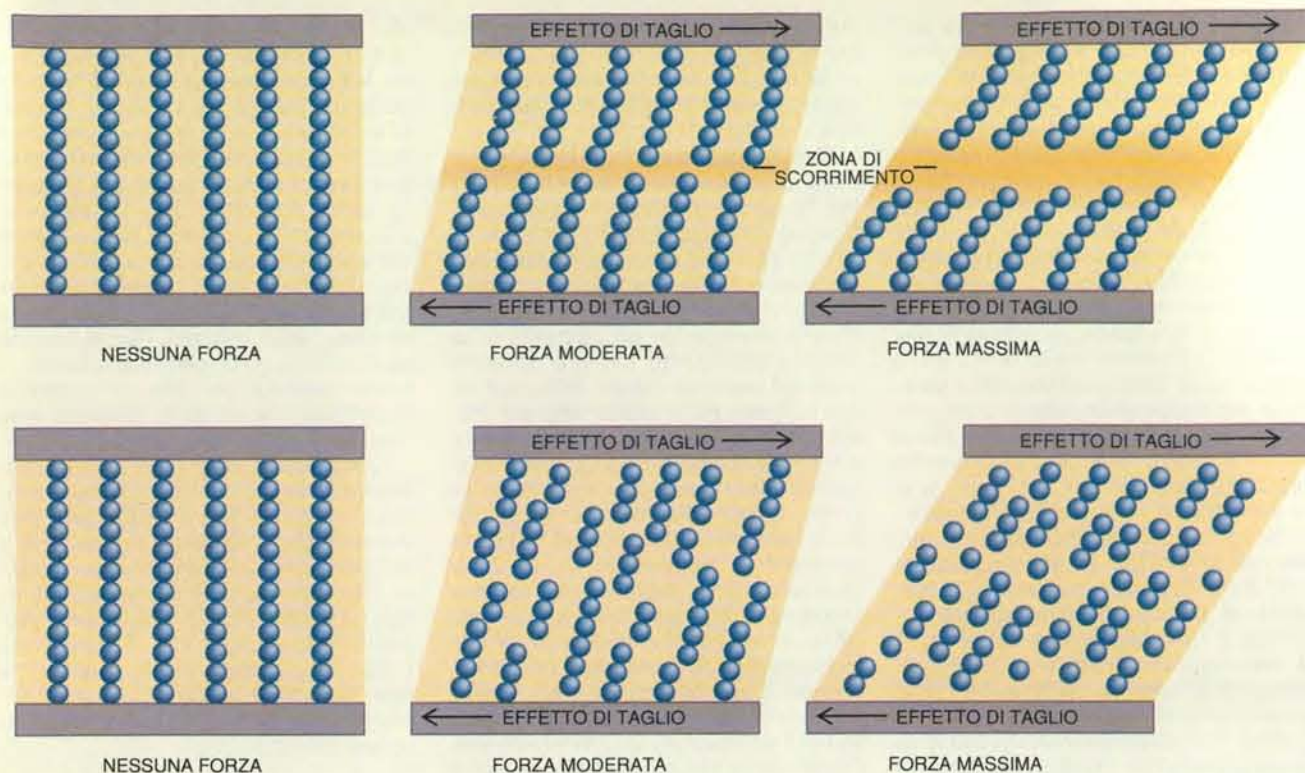
Per determinarne la struttura interna, i ricercatori hanno sfruttato le proprietà di focalizzazione delle sferette di vetro, trattando le colonne come pile di lenti capaci di trasmettere la luce. Il fascio luminoso si propagava da una sferetta all'altra fino a uscire dalla colonna. L'ultimo strato di sferette funzionava come un insieme di sorgenti luminose e creava una figura di punti sovrapposti. L'analisi del gruppo della Southern Illinois University ha permesso di dedurre che la struttura ordinata del fluido elettroreologico solidificato corrisponde a un reticolo tetragonale a corpo centrato, nel quale effettivamente le particelle di catene adiacenti sono scalate.

Sebbene le forze di attrazione fra catene siano abbastanza intense da creare il reticolo, sono però troppo deboli per spiegare l'aggregazione relativamente rapida (della durata di pochi minuti) osservata nell'esperimento. Evidentemente la realtà è più complessa di quanto faccia pensare un'analisi ordinata e precisa delle forze elettriche.

Negli esperimenti le colonne si formano più velocemente del previsto perché le catene contengono imperfezioni. Vi sono almeno due motivi per cui le catene di particelle non sono strutture perfettamente diritte e ordinate: il primo è che a volte esse non riescono a coprire l'intera distanza fra gli elettrodi; il secondo è che la loro forma è dettata in parte dal moto browniano delle particelle. Osservato per la prima volta nel 1827 dal botanico scozzese Robert Brown, questo fenomeno è un incessante moto casuale delle particelle sospese in un fluido, dovuto alle collisioni fra le particelle stesse e le molecole del fluido. Le particelle in sospensione di un fluido elettroreologico danzano caoticamente intorno alla loro posizione media nella catena a causa delle collisioni con le molecole del liquido provenienti da ogni direzione. Così, sebbene la catena possa essere in media diritta, in ogni dato momento è ondulata per il moto browniano. Questa lieve deformazione accresce la forza di attrazione fra le catene e ne catalizza l'aggregazione in colonne.

Per comprendere meglio l'accrescimento delle colonne abbiamo impiegato, in collaborazione con Judy Odinek dei Sandia National Laboratories, un metodo ottico molto diverso dalla tecnica di Chen, Zitter e Tao. Anziché dirigere il fascio luminoso lungo catene costituite da sferette di vetro relativamente grandi, abbiamo puntato un fascio laser perpendicolarmente alle catene e utilizzato sferette molto più piccole, di diametro inferiore a un nanometro. La luce diffusa da colonne formate da queste particelle contiene informazioni relative alla spaziatura e allo spessore delle colonne. In generale, la luce uscente produce una figura a forma di otto che inizialmente è larga e diffusa; via via che le colonne si ispessiscono e la distanza fra esse aumenta, la luce che si propaga nel fluido incontra meno ostacoli, la diffusione diminuisce e i lobi della figura diventano più piccoli e luminosi.

Questa regola empirica ci permette di interpretare le figure luminose osservate nei nostri esperimenti. In assenza di campo elettrico, e quindi di catene e colonne nel campione, la luce è diffusa uniformemente, ma quando si applica il campo essa comincia a concentrarsi in due lobi. Col passare del tempo il centro dell'otto diventa sempre più luminoso, un fenomeno che indica come la diffusione diminuisca via via che le colonne si ispessiscono e lo spazio libero fra di esse si allarga. All'inizio l'accrescimento è molto rapido, ma dopo pochi minuti in pratica si arresta perché la forza di attrazione fra le colonne è troppo debole per sostenerlo. Dato che secondo i calcoli il fluido in via di solidificazione potrebbe diventare più resistente con l'ispessirsi delle colonne, queste misurazioni dimostrano che mantenendo il campo elettrico per parecchi minuti si potrebbero migliorare le proprietà meccaniche dei fluidi elettroreologici.

Questi fluidi, dopo tutto, devono resistere a forze meccaniche. Conoscerne la struttura statica non è sufficiente; dobbiamo anche comprendere come essa si comporta quando viene strizzata o stirata. Questi problemi ricadono nell'ambito della reologia, che è lo studio delle proprietà di flusso dei materiali. (Quasi tutti i materiali, anche i solidi, scorrono se sono sottoposti a una pressione sufficiente, sebbene a volte questo moto sia impercettibile.)

Le figure prodotte da un fascio di luce laser che attraversa un fluido elettroreologico solidificato consentono di studiare l'organizzazione delle particelle del fluido. In questo esperimento ogni particella, fatta di vetro e molto più grande della lunghezza d'onda della luce, funge da lente. La figura che si ottiene rivela la formazione di un reticolo tetragonale a corpo centrato (*qui sopra*).

Una figura a forma di otto appare quando la luce attraversa un fluido elettroreologico le cui particelle hanno all'incirca la dimensione della lunghezza d'onda della luce. La figura indica che le particelle si sono organizzate in colonne; la sua ampiezza diminuisce all'aumentare della distanza fra le colonne.

Due teorie spiegano come le catene di particelle ostacolano lo scorrimento dei fluidi elettroreologici. Secondo quella di Klingenberg e Zukoski (*in alto*), le catene si deformano perché le loro estremità aderiscono agli elettrodi in movimento; esse si spezzano poi nel centro del fluido e la «zona di scorrimento» si allarga all'aumentare della forza di taglio. Secondo la teoria di Halsey, Martin e Adolf (*in basso*), le catene «galleggiano» nel fluido e si frammentano al crescere della forza di taglio.

La velocità di flusso di un materiale è influenzata dalla forza, o sollecitazione, a esso applicata. Per fluidi come l'acqua la velocità di flusso è proporzionale alla sollecitazione: lo scorrimento è più veloce se il pompaggio è più intenso. D'altra parte la velocità di flusso è inversamente proporzionale alla viscosità del materiale. I fluidi viscosi, come un olio denso, scorrono molto più lentamente dell'acqua a parità di forza applicata.

Il comportamento di un fluido elettroreologico immerso in un campo elettrico non è così semplice e assomiglia a quello di una classe di materiali denominati fluidi di Bingham. Al di sotto di una certa soglia di sollecitazione, un fluido di Bingham non scorre affatto, mentre al di là della soglia la velocità di flusso aumenta in proporzione alla differenza fra la forza applicata e il valore di soglia. Un esempio comunissimo di fluido di Bingham è la pasta dentifricia. Le sue proprietà reologiche permettono alla pasta strizzata dal tubetto di mantenere una forma cilindrica sullo spazzolino; allo stesso modo, i fluidi elettroreologici resistono a sollecitazioni di intensità limitata, ma cominciano a fluire quando si applica una forza maggiore.

Spesso in un fluido elettroreologico si accumula sollecitazione quando gli elettrodi cominciano a muoversi in direzioni parallele, ma opposte. La forza che così si crea è detta di taglio. Se la forza applicata è piccola, le catene di particelle del fluido si piegano e si stirano, ma non si spezzano. Tuttavia, quando le catene sono piegate, la forza di attrazione fra le particelle diminuisce perché viene meno il loro allineamento strettamente parallelo al campo; la configurazione incurvata le porta in una situazione più simile a quella in cui le particelle si trovano fianco a fianco. Se le catene si piegano troppo, le particelle escono dalla zona di mutua attrazione, le catene si spezzano e il materiale comincia a scorrere. La quantità di forza applicata che produce questo fenomeno è detta limite di snervamento.

In applicazioni quali la frizione elettroreologica, è molto importante fare in modo che il limite di snervamento sia il più alto possibile. Un fluido con un limite di snervamento elevato può resistere a una torsione considerevole prima che abbia inizio lo scorrimento e la frizione scivoli. Un modo per aumentare la resistenza di un fluido alle sollecitazioni è quello di applicare una tensione elettrica più elevata. La quantità di separazione fra cariche positive e negative sulle particelle dipolari aumenta in proporzione alla tensione applicata e la forza di attrazione fra particelle è direttamente proporzionale al quadrato della separazione di carica. Il limite di snervamento è quindi proporzionale al quadrato della tensione applicata. L'aumento della tensione, tuttavia, accresce anche il consumo di energia e riscalda il fluido elettroreologico. Dato che questo può decomporsi se diventa troppo caldo, uno degli obiettivi più importanti nella messa a punto di fluidi elettroreologici resistenti è quello di minimizzare la tensione necessaria per ottenere un dato limite di snervamento.

La tensione richiesta è in gran parte dipendente dalla polarizzabilità delle particelle. Alcuni anni fa si riteneva che la resistenza di un fluido solidificato non aumentasse in modo rilevante con la polarizzabilità, ma studi su particelle ceramiche altamente polarizzabili condotti da Douglas B. Adolf, Robert A. Anderson, Terry J. Garino e Bradley G. Hance dei Sandia National Laboratories hanno dimostrato che le cose non stanno così. I quattro ricercatori hanno osservato che la resistenza cresce in maniera evidentissima all'aumentare della polarizzazione. Vi sono quindi buone ragioni per pensare che, usando particelle altamente polarizzabili, si possano produrre fluidi elettroreologici di elevata resistenza.

La conduttività elettrica è un altro punto da tenere presente. Le differenze di conduttività fra particelle e solvente liquido sono particolarmente significative se il campo elettrico è applicato per un tempo superiore a un centesimo di secondo circa. Si considerino particelle non conduttrici in sospensione in un liquido debolmente conduttore. Quando si applica il campo, le particelle si polarizzano immediatamente, ma attraverso il liquido cominciano anche a scorrere cariche che si accumulano alla superficie delle particelle polarizzate. Dopo breve tempo, le interazioni fra particelle saranno determinate non tanto dalla loro polarizzazione intrinseca, quanto dalle nubi di carica che circondano ciascuna particella. La forza che si esercita fra le particelle è ancora essenzialmente dipolare, ma la sua origine a livello microscopico è differente, e diverso può essere anche il limite di snervamento.

Altri problemi insorgono nel caso si abbiano particelle conduttrici in un liquido isolante. Quando le particelle si polarizzano, gli elettroni non solo si spostano all'interno delle molecole, ma migrano addirittura da molecola a molecola lungo tutta la particella, creando una intensa polarizzazione e rilevanti forze attrattive. Ma se due particelle entrano in contatto reciproco, la carica fluisce dall'una all'altra; in tal caso, una delle particelle finirà per avere una carica negativa e l'altra una carica positiva. Invece di legarsi l'una all'altra, queste particelle migreranno verso elettrodi opposti, dato che questi esercitano un'attrazione assai più intensa di quella di una piccola particella. Questo fenomeno è stato osservato in sospensioni di particelle conduttrici e viene chiamato colloquialmente effetto della «palla che rimbalza». Lo scambio di cariche fra particelle conduttrici può essere evitato rivestendo queste ultime con un sottile strato isolante; il rivestimento consente alle particelle di assumere una polarizzazione elevata, e il fluido può quindi presentare un effetto elettroreologico rilevante.

Un limite di snervamento elevato non è però la sola proprietà utile di un fluido elettroreologico. Per certe applicazioni, il fluido dovrebbe cedere a una debole sollecitazione, come fa la pasta dentifricia, ma continuare a resistere al moto via via che la forza applicata aumenta. Per controllare la densità variabile di un fluido di questo tipo è necessario approfondire la conoscenza del comportamento dei fluidi elettroreologici dopo che hanno ceduto e cominciano a scorrere. Insieme con diversi altri gruppi di ricerca esaminiamo perciò quali tipi di strutture si evolvano quando un fluido è sottoposto a sollecitazione, cercando di stabilire come queste strutture inibiscano il flusso del liquido nel quale si trovano e come questo flusso a sua volta distorca e indebolisca le strutture solide in modo che il liquido continui a scorrere.

Finora non esiste alcun inquadramento teorico generale. Alcuni ricercatori hanno proposto un modello di flusso elettroreologico basato sull'osservazione di un fluido a base di amido di mais, mentre il nostro gruppo, sperimentando altri fluidi, ha trovato una spiegazione diversa del loro moto denso e lento.

Daniel J. Klingenberg, ora all'Università del Wisconsin, e Charles F. Zukoski dell'Università dell'Illinois hanno studiato il comportamento di un fluido elettroreologico, a base di amido di mais, sottoposto a sollecitazioni oltre il limite di snervamento. Essi hanno osservato che muovendo gli elettrodi in direzioni parallele e in senso opposto si ottenevano due tipi di strutture. Presso le lastre si formavano colonne inclinate di particelle, mentre nella regione intermedia si evidenziava una «zona di scorrimento» nella quale le catene erano spezzate e il fluido si muoveva liberamente. Al crescere della velocità di spostamento degli elettrodi, la zona di scorrimento aumentava la propria ampiezza e «lubrificava» il moto relativo delle colonne inclinate aderenti alle lastre.

In collaborazione con Adolf, abbiamo proposto un'interpretazione alternativa dell'azione delle forze di taglio nei fluidi elettroreologici, nella quale l'adesione delle colonne agli elettrodi non è predominante. Se le colonne sono libere di andare alla deriva, le forze di attrazione fra particelle hanno l'effetto di allineare le colonne perpendicolarmente agli elettrodi, mentre il moto del fluido tende a farle ruotare in modo che siano parallele agli elettrodi e alla direzione del flusso. Le forze generate dal flusso sono dette forze viscose. A certi angoli di inclinazione delle colonne, le forze viscose che tentano di disporre le catene parallelamente al flusso saranno compensate dalle forze elettriche che tendono a mantenerle perpendicolari. Prevediamo quindi di osservare colonne inclinate in tutto il fluido quando questo cede alle forze di taglio e comincia a scorrere.

Ci aspettiamo anche che le colonne o le catene si accorcino all'aumentare della velocità di spostamento degli elettrodi. Dato che una catena lunga ha un'area superficiale più grande di una corta, è anche maggiormente soggetta all'azione delle forze viscose. Effettivamente forze viscose intense riescono a far ruotare una catena molto lunga fino a portarla al di là dell'angolo di mutua attrazione fra particelle; a questo punto le particelle cominciano a respingersi e le catene si spezzano. Questo fenomeno fissa un limite superiore alla lunghezza delle catene che possono apparire per una data forza di taglio in un dato campo elettrico. Anche catene relativamente corte fanno tuttavia aumentare in modo rilevante la viscosità del fluido. La situazione è analoga a quella di pompare acqua attraverso una tubatura riempita di aghi di pino; il fatto che l'acqua debba farsi strada in spazi ridotti fa aumentare la forza necessaria per ottenere una certa velocità di flusso: l'acqua mescolata con aghi di pino ha una viscosità elevata, proprio come accade ai fluidi elettroreologici al di là del limite di snervamento.

Il comportamento di molti fluidi elet-

Quando a un fluido disordinato (*in alto*) si applica un campo elettrico, la solidificazione avviene in fasi successive. In qualche millisecondo si formano lunghe catene (*al centro*) che alcuni secondi dopo iniziano ad aggregarsi in colonne (*in basso*). Le particelle di silice delle fotografie hanno diametro di un nanometro.

Un materiale adattativo straordinario è mostrato nel film *Terminator 2*; qui il Terminator si mimetizza nel pavimento a piastrelle, trasformandosi rapidamente da solido a liquido. I fluidi elettroreologici reali sono ovviamente assai meno sofisticati.

troreologici può forse essere spiegato combinando il modello di Klingenberg e la nostra teoria. Presso gli elettrodi dovrebbe formarsi una regione relativamente quiescente simile al reticolo del fluido elettroreologico solidificato nello stato di riposo, mentre nella regione intermedia la struttura dovrebbe spezzarsi in brevi catene inclinate di particelle.

Oggi i fluidi elettroreologici non sono più le miscele abbastanza semplici degli anni quaranta; oltre al liquido e alle particelle in esso sospese, contengono spesso un agente disperdente e un attivatore. Il primo impedisce alle particelle di aderire l'una all'altra in assenza di campo elettrico; in sua mancanza le particelle talvolta si aggregano o non riescono a disperdersi adeguatamente anche quando l'intensità del campo elettrico è nulla, un comportamento questo che rende i fluidi di scarsa utilità. L'attivatore, che spesso è acqua o talvolta alcool, contiene impurezze in forma di sali disciolti. Il meccanismo d'azione dell'attivatore non è del tutto chiaro, ma si ritiene che l'acqua, respinta dal liquido di sospensione oleoso, si raccolga alla superficie delle particelle. I sali disciolti si polarizzano in presenza del campo elettrico e la loro carica si somma alla polarizzazione intrinseca delle particelle.

A parte la debolezza strutturale, queste formulazioni standard presentano in generale due problemi. In primo luogo, poiché l'intensità dell'effetto elettroreologico aumenta con le dimensioni delle particelle, si preferisce utilizzare particelle grandi - di diametro molto maggiore di un nanometro - che però tendono ad affondare nel fluido. (Inoltre i fluidi che contengono particelle così grandi sono molto più abrasivi che non le sospensioni di particelle più piccole.) Il secondo problema è che un attivatore come l'acqua tende a evaporare o a subire reazioni chimiche con gli altri ingredienti, abbreviando così la vita operativa del fluido, soprattutto ad alte temperature.

I fluidi elettroreologici della prima generazione, messi a punto da Winslow, dovevano essere esposti ad aria umida per conservare il proprio tenore d'acqua e contenevano particelle molto irregolari di silice. Circa 10 anni fa James E. Stangroom e un gruppo del Cranfield Institute of Technology in Inghilterra idearono una seconda generazione di fluidi con particelle meno abrasive costituite da polimeri, ma ancora contenenti acqua. Lo stadio successivo, a cui si sta lavorando attualmente, consiste nel creare una terza generazione di fluidi privi di acqua e con particelle piccole e altamente polarizzabili.

Nel frattempo anche le macchine che potrebbero impiegare dispositivi elettroreologici sono cambiate rispetto a quanto si prevedesse negli anni quaranta. L'industria automobilistica è tuttora interessata a materiali adattativi con tempi di risposta dell'ordine del millisecondo, ma questa tecnologia potrebbe essere sfruttata da un'altra industria, che mezzo secolo fa non esisteva. Oggi gran parte delle operazioni di produzione industriale sono svolte da semplici robot. Perché le recenti innovazioni nella tecnologia dei semiconduttori possano condurre alla rivoluzione robotica immaginata da generazioni di scrittori di fantascienza, saranno necessari materiali adattativi che colleghino microprocessori veloci a controlli meccanici ugualmente veloci. Un robot abbastanza agile da eseguire movimenti che per un essere umano sono del tutto usuali - per esempio afferrare una pallina al volo - potrebbe richiedere un sistema idraulico che sfrutti la risposta rapida dei fluidi elettroreologici.

Non abbiamo certo bisogno di robot per sostituire gli atleti umani, ma si immagini per esempio un dispositivo controllato da calcolatore che estruda i finissimi cavi metallici usati per collegare i circuiti integrati. Per avvolgere su una spoletta questi cavi delicatissimi il robot dovrebbe controllare la tensione con estrema esattezza: il cavo potrebbe allentarsi o spezzarsi se la macchina non riuscisse ad avvertire rapidamente e a correggere anche il minimo cambiamento di tensione. Un prototipo di un dispositivo di questo genere, dotato di una frizione elettroreologica veloce, è già stato sperimentato, e altri potrebbero ben presto seguirne le orme: queste macchine avrebbero a tutti gli effetti «articolazioni» sovrumane capaci di chiudersi o allentarsi istantaneamente con precisione estremamente elevata. Naturalmente non potrebbero tener testa al Terminator T-1000, ma sarebbero altamente competitive nel mercato della tecnologia.

BIBLIOGRAFIA

WINSLOW W. M., *Induced Fibration of Suspensions* in «Journal of Applied Physics», 20, n. 12, dicembre 1949.

GAST ALICE P. e ZUKOSKI C. F., *Electrorheological Fluids as Colloidal Suspensions* in «Advances in Colloid and Interface Science», 30, n. 3-4, dicembre 1989.

HALSEY THOMAS C., *Electrorheological Fluids* in «Science», 258, pp. 761-766, 30 ottobre 1992.

La modularità delle proteine nell'evoluzione

Moltissime proteine sono costituite da un insieme limitato di elementi modulari che si sono diffusi e moltiplicati nel corso dell'evoluzione per mezzo di meccanismi che solo oggi cominciano a essere definiti

di Russell F. Doolittle e Peer Bork

Negli ultimi decenni gli studiosi di biologia molecolare e di biochimica hanno scoperto che parecchie proteine constano di dominî, ossia blocchi distinti di amminoacidi, molti dei quali hanno funzioni ben definite che contribuiscono all'attività globale della molecola. Spesso, inoltre, nel corso dell'evoluzione alcune di queste unità modulari si spostano all'interno della proteina o da una proteina all'altra. Questa mobilità evolutiva non è limitata all'ambito della specie; in alcuni casi i dominî sono in grado di spostarsi dalla linea di discendenza di una specie a quella di un'altra, passando addirittura da cellule animali a cellule batteriche.

Dato che anche nei segmenti genici che codificano per le proteine è stata rilevata una divisione in compartimenti, molti biologi ritengono che questi aspetti strutturali altro non siano che il riflesso di uno stesso fenomeno di base; ogni regione codificante di un gene corrisponderebbe a una caratteristica strutturale specifica di una proteina. Noi e i nostri collaboratori abbiamo un'opinione leggermente diversa; a nostro parere, gli elementi disponibili indicano che la suddivisione dei geni in parti codificanti separate è un evento di gran lunga più recente da un punto di vista evolutivo.

Le proteine sono lunghe catene formate da molecole di piccole dimensioni, gli amminoacidi. Venti diversi amminoacidi, ciascuno con forma e carattere chimico proprio, costituiscono tutte le proteine esistenti in natura. Le proprietà di una proteina dipendono da quali dei 20 amminoacidi entrino a far parte della molecola e soprattutto dal loro ordine sequenziale. In particolare, è proprio la sequenza degli amminoacidi a determinare il modo in cui la proteina si avvolge per dare origine a una molecola tridimensionale dotata di attività.

La lunghezza di una sequenza è un fattore importante per l'avvolgimento. Le catene polipeptidiche possono essere costituite da molte migliaia di amminoacidi. (Finora il primato appartiene alla titina, una proteina dei muscoli che contiene fino a 30 000 residui amminoacidici.) L'estensione delle catene più corte, però, non permette un numero di legami intramolecolari tale da dare luogo a un'unica conformazione: esse tendono quindi a oscillare da una forma all'altra. Quando una catena polipeptidica contiene 30 o 40 residui, comincia ad avere forze di coesione sufficienti per assumere una forma predominante, anche se può occorrere un'ulteriore stabilizzazione data da legami con ioni metallici o da ponti disolfuro come quelli che si instaurano tra due residui di cisteina.

In ambiente costante, qualunque proteina che contenga un numero di unità amminoacidiche superiore a un minimo si avvolgerà sempre allo stesso modo. Questo ambiente può essere o la soluzione salina che costituisce molti liquidi biologici o il materiale lipidico delle membrane biologiche; può anche includere proteine adiacenti o addirittura altre parti della catena polipeptidica stessa.

Una sequenza che, in un insieme ben definito di circostanze, si avvolga spontaneamente in una conformazione caratteristica è detta dominio. Questa definizione formale è applicata di rado in modo rigoroso; il più delle volte il termine vale per qualsiasi parte di una proteina che possa essere considerata strutturalmente distinta dal resto. Alcune proteine di piccole dimensioni sono costituite da un unico dominio; molte altre sono composte da due o più dominî, e alcune contengono più dominî, le cui forme possono essere molto simili o molto diverse.

Il modo più diretto per identificare un dominio consiste nel determinarne la struttura mediante diffrazione di raggi X da parte dei cristalli della proteina oppure mediante studi di risonanza magnetica nucleare (NMR). Una volta identificata la sequenza amminoacidica di un dominio, si possono trovare altri dominî affini senza dover ricorrere a indagini strutturali: basta semplicemente cercare sequenze amminoacidiche simili a quelle delle strutture note. Questa scorciatoia è estremamente utile perché le sequenze amminoacidiche conosciute sono molto più numerose che non le strutture ricavate mediante diffrazione di raggi X o risonanza magnetica nucleare. Anche in assenza di determinazioni sperimentali è spesso possibile dedurre l'esistenza di un dominio esclusivamente in base alla sequenza. Osservando in questo modo le somiglianze strutturali e di sequenza delle proteine, si possono apprendere molte cose sulla loro evoluzione.

Fino all'inizio degli anni settanta, la concezione tradizionale dell'evoluzione delle proteine si basava principalmente su «duplicazione e modificazione». Il gene responsabile della sintesi di una data proteina viene occasionalmente duplicato tramite vari processi di ricombinazione, durante i quali filamenti di DNA si scambiano l'informazione genetica. Qualche volta la duplicazione dà luogo a un secondo gene, che può subire un'ulteriore modificazione o mutazione e produrre una proteina inedita con una nuova funzione. In alternativa, il DNA può essere duplicato in tandem, nel qual caso la molecola proteica codificata viene allungata e può assumere nuove proprietà. Come appare ovvio dai confronti tra sequenze amminoacidiche, questi fenomeni di duplicazione interna hanno originato numerose proteine con molecola estesa, da quelle di piccole dimensioni come le ferredossine batteriche, che constano di soli 56 residui amminoacidici, a quelle molto grandi, come la beta-galattosidasi batterica, costituita da oltre 1000 residui amminoacidici.

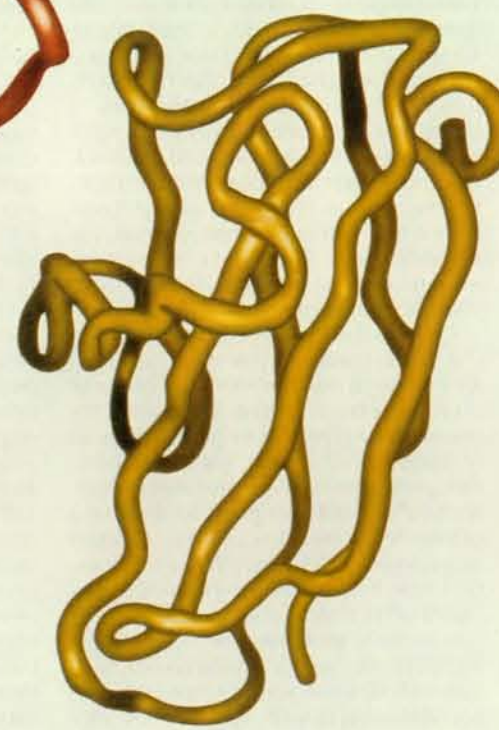
Un aspetto nascosto dell'evoluzione delle proteine venne però alla luce circa 20 anni fa, quando Michael G. Rossmann della Purdue University determinò la struttura tridimensionale della lattico-deidrogenasi mediante diffrazione di raggi X. Rossmann notò che una parte della molecola assomigliava moltissimo a certe strutture di altre proteine che egli aveva osservato in precedenza: in particolare, una regione dell'enzima che si legava a un cofattore aveva evidenti controparti in altre deidrogenasi.

Ciò che rendeva degne di rilievo queste somiglianze strutturali era il fatto che esse non comparissero sempre nelle stesse parti relative delle molecole. Pareva che, durante l'evoluzione, l'unità si fosse spostata all'interno della sequenza amminoacidica lineare senza perdere la propria funzione di legame con un cofattore. Rossmann propose che le proteine fossero costituite da moduli (che noi oggi considereremmo dominî), comparsi precocemente nella storia della vita e assemblatisi in differenti combinazioni.

Il modello di evoluzione delle proteine da lui prospettato aumentava enormemente le possibilità realizzabili per semplice duplicazione e moltiplicazione. Se nuove proteine potevano essere prodotte ricombinando componenti di proteine già esistenti, la diversità proteica sarebbe aumentata in modo esplosivo.

L'ipotesi di Rossmann si è rivelata profetica. Sono state determinate le sequenze amminoacidiche di numerose grandi molecole proteiche e si è visto che molte di esse hanno quel tipo di struttura grossolanamente ripetitiva che ci si potrebbe aspettare da una sequenza di moduli mobili. Per esempio, la fibro-

DOMINIO 2 CD2

DOMINIO 1 PapD

DOMINIO 1 GHR

Fn3 (DOMINIO 10)

In molte proteine sono state trovate sequenze geniche mobili; vengono qui illustrati due tipi di questi moduli, o dominî. Fn3 e GHR sono esempi di moduli della fibronectina di tipo III, mentre PapD e CD2 sono dominî immunoglobulinici. Si tratta di sequenze lineari di amminoacidi che possono avvolgersi in strutture costanti, riconoscibili, dotate di proprietà biochimiche specifiche. Nel corso dell'evoluzione questi dominî possono trasferirsi da una proteina all'altra come unità discrete, il che favorisce la comparsa di nuovi tipi di proteine.

Come le perle di una collana, i dominî si presentano quali subunità modulari proteiche distinte. Alcune proteine, come la fibronectina, il collagene XII e la titina, una proteina gigante dei muscoli, contengono molte copie degli stessi dominî.

nectina consta di due lunghe catene, ciascuna delle quali contiene oltre 2000 residui amminoacidici. Un primo esame ha rivelato che le catene di questa grande proteina extracellulare sono costituite da molte serie di tre tipi differenti di sequenze ripetitive. Queste, indicate come Fn1, Fn2 e Fn3, hanno rispettivamente lunghezze dell'ordine di 45, 60 e 100 residui amminoacidici. (Le sequenze ripetitive sono «imperfette», nel senso che quelle di un dato tipo non sono tutte identiche.) Presumibilmente ogni tipo di sequenza ripetitiva può avvolgersi in maniera indipendente, come un vero e proprio dominio, e la proteina nel suo complesso deve assomigliare a una lunga collana costituita da tre tipi di perle.

Sorprende il fatto che sequenze simili a Fn1, Fn2 e Fn3 siano state osservate successivamente in un numero enorme di altre proteine animali. Lo stesso vale anche per numerosi altri dominî identificati. La piccola proteina denominata fattore di crescita epidermico consiste in un unico dominio (nell'uomo, di 53 residui amminoacidici), strettamente avvolto su se stesso e fissato in questa configurazione da tre ponti disolfuro. Analoghi dominî sono stati identificati in più di 100 proteine, all'interno delle quali essi appaiono anche in oltre 30 copie.

Le funzioni di molti di questi moduli non sono ancora del tutto chiare, ma parecchi di essi si legano a particolari sostanze, o almeno le riconoscono. Vi è una famiglia di lectine che si lega a vari carboidrati. Così pure, ben noto per le sue capacità di legame è il dominio immunoglobulinico, una regione caratteristica degli anticorpi e di altre molecole del sistema immunitario. Alcuni dominî possono servire da «etichette» che identificano una proteina quale componente di un determinato tessuto. Molti di essi sembrano puri e semplici elementi di distanziamento o di connessione, mentre altri ancora non hanno probabilmente alcuna funzione. Sembra quindi che, durante l'evoluzione, molti dominî possano spostarsi all'interno di una proteina o tra una proteina e l'altra. Fino a quando non vi è danno o perdita di funzionalità in seguito al rimaneggiamento, il costo evolutivo del mantenimento di un dominio in una nuova collocazione è insignificante. Questo risultato è un corollario della teoria dell'evoluzione neutrale, ma potrebbe essere formulato come una regola mutuata dal basket professionistico: «Nessun danno, nessun fallo».

Quando Rossmann propose per primo che unità modulari potessero spostarsi all'interno di una proteina o tra proteine diverse, nessuno rifletté seriamente sui meccanismi genetici che potevano essere responsabili di tali riarrangiamenti. Ben presto, però, ci si imbatté in un'inaspettata proprietà dei geni che sembrava offrire una spiegazione. Come avevano scoperto nel 1953 James Watson e Francis Crick, l'informazione genetica è contenuta nella doppia elica del DNA. Proprio come le proteine sono costituite da sequenze di residui amminoacidici, le molecole di DNA sono sequenze di basi nucleotidiche, le quali vengono trascritte come filamenti complementari di RNA messaggero. Nelle cellule, i ribosomi assemblano le molecole proteiche secondo le istruzioni contenute nell'RNA: ogni codone, o sequenza di tre basi dell'RNA, corrisponde a un amminoacido.

Il fatto sorprendente osservato a metà degli anni settanta è che il DNA che codifica per un polipeptide può essere interrotto da sequenze non codificanti. Queste vengono escisse mediante un meccanismo di «taglia-e-cuci» prima che l'RNA messaggero sia tradotto in un polipeptide.

La scoperta di queste interruzioni nei geni ha spinto Walter Gilbert della Harvard University a ipotizzare che le sequenze non codificanti, da lui chiamate introni (sequenze intercalari), facilitassero lo scambio delle parti codificanti dei geni, definite esoni (sequenze che vengono espresse). Gilbert riteneva che la distanza addizionale tra segmenti codificanti offrisse un numero proporzionalmente superiore di possibilità di ricombinazione, dato che questa dipende da una rottura casuale della molecola di DNA. Si è avanzata l'ipotesi che una somiglianza tra sequenze di introni potesse promuovere, durante la ricombinazione, disallineamenti e scambi ineguali di DNA che faciliterebbero il riarrangiamento genico. A quell'epoca non vi era motivo per supporre che gli introni avessero sequenze simili, ma è stato poi dimostrato che in effetti essi sono un vero rifugio per molti elementi genetici mobili. Le sequenze simili di questi elementi possono contribuire agli errori genetici che si hanno durante la meiosi (il processo di divisione cellulare che dà origine a cellule uovo e spermatozoi).

In molti esseri viventi non avviene meiosi: è possibile che questi organismi si siano lasciati sfuggire un modo splendido per costruire nuove proteine? Gli introni che interrompono il codice per la sintesi delle proteine si trovano soltanto nel DNA degli eucarioti. I geni per le proteine batteriche non contengono introni; ogni gruppo di tre basi nell'acido nucleico corrisponde a un amminoacido nella proteina. (In realtà alcuni tipi di introni che non interrompono le sequenze codificanti per le proteine sono stati trovati anche nei batteri, ma non riguardano da vicino questa trattazione.)

L'assenza di introni dai geni per le proteine batteriche ha indotto Ford Doolittle della Dalhousie University e James Darnell della Rockefeller University a ipotizzare indipendentemente che in passato i batteri possedessero introni, ma li abbiano perduti. Presumibilmente il loro genoma si è semplificato nel corso dell'evoluzione per rendere più efficiente la duplicazione. Per riassumere, gli introni esistono fin dall'origine della vita, mentre le sequenze codificanti corte sono state generate separatamente.

L'ipotesi di Doolittle e Darnell ha dato luogo a una controversia che non si è ancora risolta: gli introni sono comparsi «precocemente» e sono fondamentali per l'origine di tutte le proteine, oppure hanno fatto un'apparizione «tardiva»? La seconda alternativa è stata proposta da Thomas Cavalier-Smith, oggi all'Università della British Columbia. Secondo questo autore, gli introni potrebbero essere frammenti invasivi di acidi nucleici, i cosiddetti elementi trasponibili, originatisi in organismi simbionti che hanno poi finito per trasformarsi nei mitocondri e negli altri organelli delle cellule eucariote. L'idea di Cavalier-Smith è stata poi sviluppata da altri, tra cui Donal Hickey dell'Università di Ottawa.

Si è visto che i segmenti di DNA che codificano per i moduli mobili delle proteine sono spesso, ma non sempre, fiancheggiati da introni; in altri termini, le unità strutturali di molte proteine sono codificate da esoni. Questa osservazione ha alimentato la convinzione che tutti gli esoni fossero evolutivamente mobili e che corrispondessero a possibili elementi costruttivi modulari delle proteine.

A nostro parere, un'idea del genere è errata per due motivi. In primo luogo, come ha fatto notare per primo Lázló Patthy dell'Istituto di enzimologia di Budapest, tutti gli esoni possono essere rimescolati, ma solo una frazione delle unità così spostate sarà geneticamente compatibile: ciò significa che in certe nuove posizioni molte di queste unità non potranno essere tradotte in modo sensato. Quando un introne viene a trovarsi all'interno di una sequenza codificante, occuperà uno di tre tipi possibili di posizione: esattamente tra due codoni (tipo 0), tra la prima e la seconda posizione di un codone (tipo 1) oppure tra la seconda e la terza posizione (tipo 2). Se esso e la sequenza codificante adiacente vengono spostati in un sito nuovo, l'introne dovrà avere lo stesso tipo di posizione, altrimenti i codoni fatti slittare saranno tradotti in modo inesatto e ne risulterà una sequenza amminoacidica priva di senso. Se gli introni venissero distribuiti a caso, ci si aspetterebbe che dopo il rimescolamento solo un terzo delle nuove combinazioni di esoni sia in fase. È curioso il fatto che la stragrande maggioranza dei geni che codificano per i moduli rimescolati più frequentemente sia fiancheggiata da introni di tipo 1.

L'altra ragione fondamentale per cui solo alcuni esoni sono evolutivamente mobili è che soltanto i dominî veri (cioè quelli che sono in grado di avvolgersi completamente e in maniera indipendente) possono sopravvivere all'interno di un ambiente proteico nuovo. Le sequenze più piccole e meno autosufficienti non sarebbero in grado di avvolgersi e perderebbero la propria identità. Inoltre, se un'unità spostata dovesse andare a finire tra due esoni che non fossero dominî veri, il prodotto del gene che riceve questa unità potrebbe anch'esso non avvolgersi in modo appropriato.

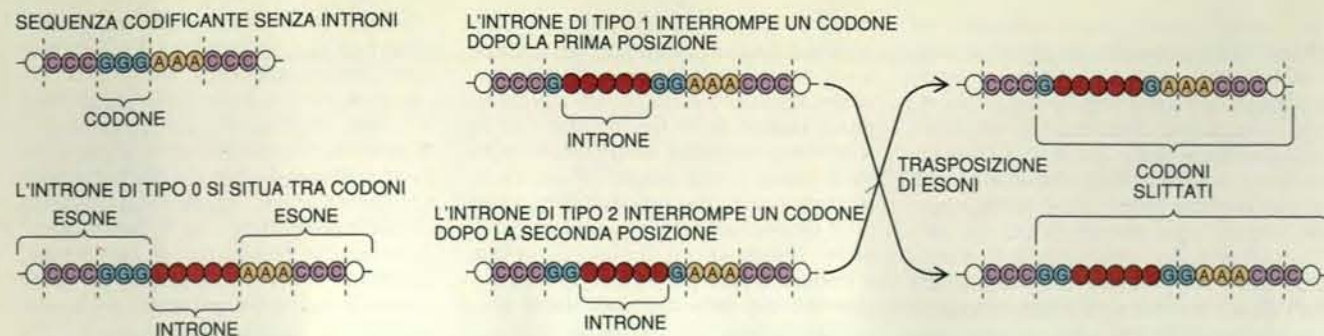
Questi due fattori, l'uno genetico e l'altro strutturale, contribuiscono a far sì che i dominî mobili si trovino spesso associati. Non di rado le proteine che contengono un dominio mobile ne contengono altri; alcune sono mosaici che includono fino a cinque differenti moduli di questo tipo. Queste proteine tollerano bene il riarrangiamento, sia sotto il profilo genetico sia sotto quello strutturale.

L'osservazione che molti moduli delle proteine sono codificati da esoni è stata interpretata come una dimostrazione del fatto che gli organismi primordiali abbiano assemblato tutte le loro proteine partendo da una scorta di componenti strutturali primitive codificate da esoni. Tuttavia, vi sono parecchi argomenti contrari a questa interpretazione. Per esempio, un semplice calcolo aritmetico dimostra che gli ipotetici esoni primitivi sarebbero stati troppo piccoli per produrre componenti proteiche in grado di avvolgersi da sole. La dimensione media degli esoni noti nel genoma dei vertebrati attuali è di 135 basi nucleotidiche, il che corrisponde a un polipeptide di appena 45 amminoacidi. Una sequenza così corta richiede generalmente una ulteriore stabilizzazione per potersi avvolgere in una conformazione stabile.

Si tenga anche presente che, secondo i sostenitori della teoria degli «introni precoci», questi vanno costantemente perduti con il trascorrere del tempo; questa conclusione è resa inevitabile dalla incostante presenza degli introni in specie differenti. Un simile andamento irregolare potrebbe essere il risultato sia dell'acquisizione sia della perdita di introni, ma, se si resta fedeli all'idea che

Dimensioni di alcuni dominî mobili

I dominî dotati di mobilità evolutiva hanno dimensioni variabili. Alcuni conservano la propria forma grazie a ponti disolfuro tra coppie di residui di cisteina, mentre altri sono stabili anche senza questi legami.

DOMINI CONTENENTI PONTI DISOLFURO	NUMERO APPROSSIMATIVO DI UNITÀ AMMINOACIDICHE	RESIDUI DI CISTEINA
Somatomedina B	40	8
Complemento C9	40	6
EGF	45	6
Fn1	45	4
Fn2	60	4
«Apple»(Sushi) (ß GP26)	61	4
Ovomucoide	68	6
VWF-C	72	8
Kringle	80	6
Kunitz	80	6
Link	~100	4
Recettore «spazzino»	110	6
Dominio correlato al fibrinogeno	250	4 o 6
NESSUN PONTE DISOLFURO		
Collagene	18	
Dominio ricco di leucina	25	
«Gla»	35	
Dominio che si lega al collagene	50	
Dominio simile alla lectina	100	
Fn3	~100	
SH2	100	
SH3	100	

La struttura interrotta di alcuni geni che codificano per le proteine fa sì che essi si dividano in esoni e introni (sequenze che vengono espresse e sequenze intercalari). Le triplette di basi nucleotidiche (codoni) degli esoni vengono tradotte in amminoacidi. Gli introni possono interrompere le sequenze codificanti in tre differenti posizioni. Se esoni e introni si spostassero a caso, le trasposizioni tra differenti tipi di introni provocherebbero mutazioni per slittamento del sistema di lettura. Questo problema rende poco probabile l'ipotesi secondo cui tutti gli esoni codificherebbero per moduli mobili.

questi siano esistiti fin dalle origini, la sola spiegazione possibile è la perdita. Di conseguenza i primi esoni sarebbero stati ancora più corti e avrebbero codificato per polipeptidi verosimilmente incapaci di avvolgersi da soli in dominî.

Un altro argomento contrario all'uso delle attuali proteine a mosaico quali esempi di unità costruttive primordiali è la distribuzione dei dominî tra le proteine. La stragrande maggioranza dei moduli mobili finora noti si trova esclusivamente nelle proteine animali. Allo stato attuale della ricerca non sappiamo realmente quando o dove la maggior parte di essi sia inizialmente comparsa. Forse le tracce della loro evoluzione sono state in parte nascoste da estesi cambiamenti di sequenze avvenuti nei dominî corrispondenti delle piante, dei funghi e dei protozoi. Come diremo, il fatto che le strutture tridimensionali siano più persistenti, in senso evolutivo, delle sequenze può aiutarci a risolvere questo enigma.

A parte tutti gli elementi secondo i quali la maggior parte degli esoni non sarebbe dotata di mobilità evolutiva, è anche chiaro che alcuni dominî mobili non sono il prodotto di singoli esoni. Un grande dominio, osservato per la prima volta nella molecola di fibrinogeno dei vertebrati, è composto da 250 amminoacidi. In alcune proteine il gene responsabile della sua sintesi contiene molteplici introni; eppure nessuno dei singoli esoni è mai stato trovato da solo, in assenza di tutti gli altri esoni. A quanto pare, nessuno degli esoni si è mai trasferito al di fuori del dominio. Così, la semplice presenza di introni in un gene non è sufficiente a rendere mobili gli esoni. Il fatto che la grande maggioranza degli esoni identificati non sia mai stata trovata in più di una collocazione depone contro una mobilità semplice e indiscriminata.

Vi sono altri esempi di unità mobili che contengono introni all'interno delle loro sequenze codificanti. Uno dei primi moduli mobili identificati è stato chiamato *kringle* (per la somiglianza di forma con un biscotto danese che porta questo nome). Esso consta di circa 80 unità amminoacidiche e contiene tre caratteristici ponti disolfuro; è molto simile al dominio Fn2, eccetto per il numero di residui amminoacidici che si trovano tra i residui di cisteina, tanto che alcuni ricercatori non fanno differenza tra i due. In due sue collocazioni, il gene per il kringle è scisso da un introne, ma finora nessuno ha trovato la metà di un kringle in una qualsiasi proteina.

Un ulteriore punto a favore della teoria degli «introni tardivi» è che gli introni che interrompono le regioni codificanti sono molto più comuni nelle piante e negli animali che non negli eucarioti diversificatisi precocemente; in eucarioti primitivi come *Giardia lamblia* non è stato trovato alcun introne. Oltre a ciò nelle piante sono state identificate proteine modulari di cui non si conosce alcuna controparte negli animali, e viceversa. Infine vi sono prove indirette del fatto che l'assemblaggio modulare di alcune proteine batteriche sia avvenuto in epoca così recente che queste proteine devono essersi necessariamente evolute senza l'aiuto di introni. Tutto ciò fa pensare che gli introni abbiano fatto la loro comparsa in epoca successiva all'evoluzione degli eucarioti.

Così alcuni esoni codificano per i dominî, ma la maggior parte di essi non lo fa; gli esoni codificanti possono spesso venire duplicati geneticamente e spostati. La possibile esistenza di una relazione di causa ed effetto tra questi fenomeni fa sorgere un problema spinoso. Forse l'evoluzione degli introni ha facilitato il rimescolamento degli esoni; d'altro canto, non è impossibile che questi introni spesso delimitino le regioni che codificano per i dominî, proprio perché questa collocazione è vantaggiosa per la loro propagazione. Se un introne interrompesse una sequenza codificante per un dominio, potrebbe sopravvivere in questa situazione (ammesso che non violi la regola della fase ricordata prima), ma non si diffonderebbe ulteriormente perché gli esoni ai suoi lati non sarebbero autosufficienti e non potrebbero muoversi in maniera indipendente l'uno dall'altro. Per contro, se l'introne si colloca tra regioni che codificano per unità in grado di avvolgersi indipendentemente, può diffondersi in altre posizioni assieme agli esoni. In questo modo il rimescolamento degli esoni può essere solo marginale per la sopravvivenza degli introni.

Un modo per saperne di più sull'evoluzione dei moduli mobili è quello di osservare in particolare la struttura e la diffusione di uno di essi. Il nostro favorito è Fn3, il dominio della fibronectina di tipo III. Come i kringle, le unità Fn3 vengono talvolta interrotte da singoli introni; tuttavia esse non sono mai state osservate con un numero di amminoacidi inferiore a 90-100, che rappresenta il loro corredo completo. Per parecchi anni ci siamo interessati, indipendentemente, alla scoperta di dominî Fn3 in varie proteine. All'inizio la sequenza sembrava apparire solo nelle proteine animali, cosicché siamo rimasti sorpresi quando, nel 1990, ricercatori della Niigata University in Giappone hanno riferito dell'esistenza di Fn3 in una proteina batterica. Il nostro comune interesse è emerso in occasione di un convegno tenutosi nel 1991 in Italia; a quel punto abbiamo deciso di unire i nostri sforzi per compilare un inventario delle presenze di Fn3.

A questo scopo abbiamo esaminato una base di dati di sequenze proteiche con vari mezzi, tra cui un algoritmo per la ricerca di configurazioni che uno di noi (Bork) aveva ideato insieme con Christian Grunwald all'Istituto centrale di biologia molecolare di Berlino Est. Abbiamo scoperto ben più di 300 presenze distinte del motivo Fn3, il che ci ha permesso di stabilire che esso era un vero dominio, se mai vi fossero stati dubbi. Le 300 presenze rappresentavano in realtà 67 proteine differenti se non si contavano proteine uguali di specie diverse. Fra queste, sessanta erano di origine animale e sette di origine batterica; nessuna delle sequenze identificate proveniva da piante, funghi o eucarioti unicellulari.

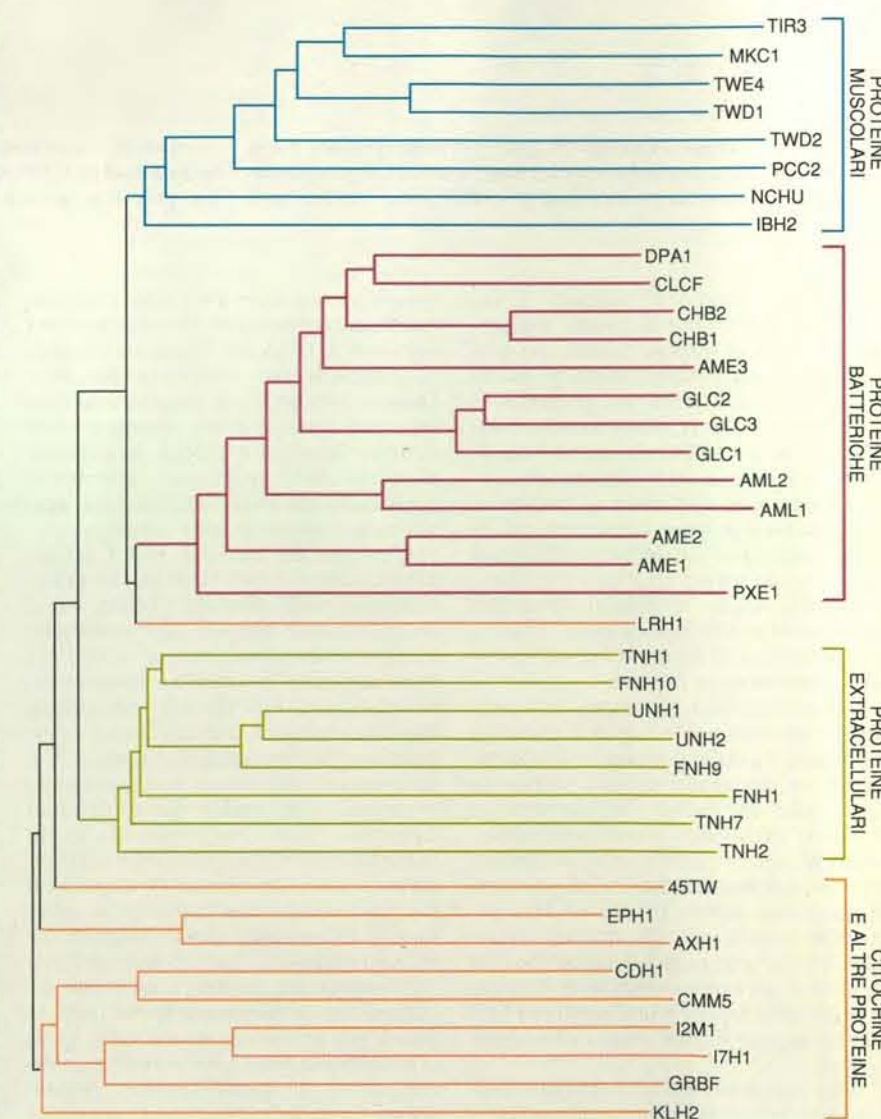
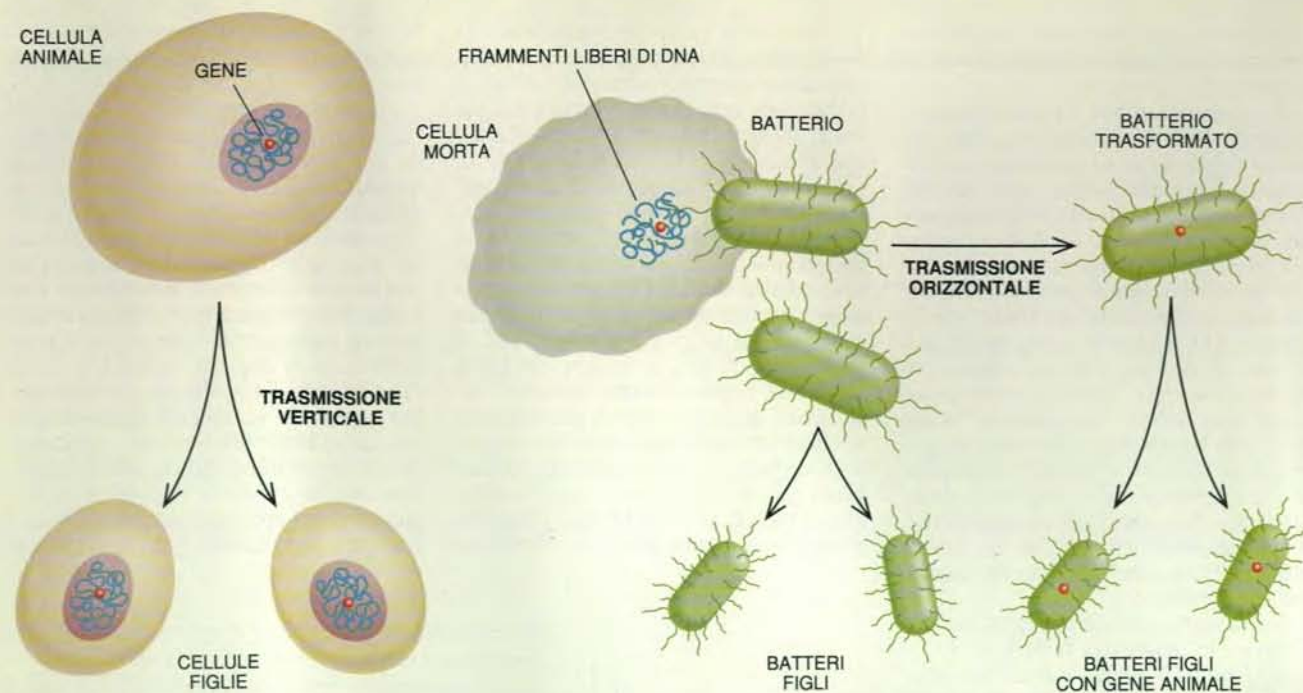
Le domande ovvie da porsi erano le seguenti: batteri e animali avevano ereditato entrambi il dominio da qualche antenato comune, oppure uno dei due gruppi lo aveva acquisito in qualche modo dall'altro gruppo? Se il dominio era già presente nell'antenato comune di procarioti ed eucarioti, perché allora non era stato trovato anche nei funghi e nelle piante? Al calcolatore, abbiamo allineato tutte le sequenze Fn3 che eravamo riusciti a trovare e abbiamo costruito un grossolano albero filogenetico basato sulla loro somiglianza. Dato che eseguire questo lavoro con tutte le 300 sequenze era impresa davvero improba, abbiamo cominciato con insiemi rappresentativi di sequenze provenienti da tutte le proteine batteriche e da quelle animali più differenziate.

È stato subito chiaro che qualcosa non andava: le sequenze batteriche erano troppo simili a quelle animali per poter essere derivate da un antenato comune di due miliardi di anni fa. Viceversa i dati disponibili (tra cui alberi filogenetici ottenuti al calcolatore) facevano pensare che i batteri avessero acquisito in qualche modo il dominio Fn3 da una fonte animale.

C'erano parecchi motivi per giungere a questa conclusione. In primo luogo accadeva spesso che un enzima di una specie batterica possedesse il dominio in questione, mentre lo stesso enzima di altre specie non lo conteneva, il che sottintendeva che quel dominio fosse strutturalmente e funzionalmente superfluo e che quindi fosse giunto tardivamente in certi batteri. Inoltre le sequenze Fn3 apparivano sia sporadicamente sia a gruppi, ma sempre in un insieme caratteristico di enzimi extracellulari; se i batteri perdessero copie di Fn3 col passare del tempo, ci si aspetterebbe di trovare le copie superstiti in una maggiore varietà di proteine.

Infine, benché i batteri che possedevano i dominî Fn3 fossero di tipi diversi, avevano certe caratteristiche in comune. Erano tutti batteri del suolo che ricavavano il nutrimento da polimeri comuni, come la chitina e la cellulosa, che si liberano per decomposizione di materiale organico. Sono stati esaminati numerosissimi altri tipi di batteri, ma nessuno presentava il dominio Fn3 in una delle sue proteine. È nota oltre metà della sequenza genomica del batterio intestinale *Escherichia coli* e non è apparsa alcuna traccia della presenza di Fn3; lo stesso si può dire per il gran numero di sequenze studiate nei funghi e nelle piante. Se il dominio Fn3 fosse stato presente in un antenato comune dei procarioti e degli eucarioti, ci si aspetterebbe che la sua radiazione fosse proseguita lungo le linee di discendenza principali, e che oggi esso fosse rappresentato in tutti questi gruppi.

L'idea che il gene che codifica per un dominio possa migrare fra organismi molto diversi potrebbe sembrare a prima vista stravagante: l'esperienza comune mostra che i geni si trasmettono «verticalmente», da una generazione di organismi alla successiva. È comunque possibile che a volte i geni siano trasmessi «orizzontalmente», non soltanto da specie a specie, ma anche tra linee di discendenza distanti. Alcuni virus possono catturare piccoli geni di un ospite e trasportarli in un altro; in casi rari, il gene trasposto può assimilarsi nel DNA del nuovo ospite. I batteri possono trasformarsi assumendo DNA dall'ambiente circostante, nel quale sono spesso presenti cellule animali in decomposizione. Molti batteri hanno anche geni localizzati su piccoli anelli di DNA, o plasmidi, che possono scambiare con altri batteri. In teoria tutti questi meccanismi rappresentano altrettante opportunità di trasmissione orizzontale dei geni.

Ammettendo che qualche batterio abbia acquisito il gene per il dominio Fn3 da una fonte animale, quando avvenne questo passaggio? Tutto ciò che l'albero filogenetico può indicare è che l'evento si è verificato nell'ultimo miliardo di anni, dopo la divergenza degli animali da una parte e delle piante e dei funghi dall'altra. Per precisare la cronologia occorrerebbe conoscere la velocità media con cui variano le sequenze lungo le linee di discendenza sia batterica sia animale. Per le proteine animali possiamo stimare questa velocità confrontando sequenze di vari organismi di cui sia nota dalla documentazione fossile l'epoca di divergenza; purtroppo non possediamo informazioni paragonabili per le sequenze

In quest'albero filogenetico sono evidenziate le relazioni evolutive di 39 dominî Fn3 provenienti da varie fonti. L'albero è stato costruito al calcolatore, confrontando le sequenze dei dominî di animali e batteri. Sorprendentemente, i dominî dei batteri sono molto affini ad alcuni dominî degli animali, e ciò fa pensare che i batteri abbiano ricevuto dagli animali i geni responsabili della sintesi di questi moduli.

La trasmissione «orizzontale» di geni potrebbe spiegare come i batteri abbiano acquisito dominî dagli animali. I geni passano verticalmente tra generazioni di cellule, ma i batteri possono anche assorbire DNA dall'ambiente. Assumendo da una cellula morta il DNA che codifica per un dominio, un batterio potrebbe quindi trasmettere questo gene alla progenie.

batteriche. (Sono stati scoperti alcuni microfossili riferibili ai batteri, ma certamente non si dispone ancora per essi di una gerarchia evolutiva interpretabile come quella che si ha per gli animali.) È possibile, tuttavia, riconoscere sia nelle proteine animali sia in quelle batteriche una tendenza delle sequenze Fn3 alla duplicazione in tandem: all'interno di una qualsiasi proteina contenente più di un dominio Fn3, le sequenze di questi dominî sono spesso adiacenti e, in genere, molto simili. Questa osservazione sottintende che la duplicazione del DNA codificante per il dominio Fn3 debba essere relativamente recente.

La collocazione nel tempo della trasmissione orizzontale e delle duplicazioni di geni ha un'importanza cruciale per capire in che modo si siano diffuse le unità genetiche mobili. Per quanto si sa, i batteri attuali non possiedono introni nelle sequenze geniche che codificano per le proteine. Se mai in passato essi hanno avuto introni all'interno delle sequenze codificanti, da quanto tempo questi sono scomparsi? A meno che non si tratti di un evento abbastanza recente, è necessario supporre che i geni per Fn3 devono essersi diffusi senza l'aiuto degli introni.

Una possibilità è che il dominio Fn3 sia stato diffuso tra i batteri del suolo da un batteriofago (virus batterico) o da un plasmidio promiscuo. Speriamo di riuscire prima o poi a cogliere la trasmissione in atto trovando un fago che contenga un gene per un dominio Fn3. Dal momento che sono state ormai trovate numerose sequenze Fn3 batteriche, dovrebbe essere possibile sintetizzare brevi segmenti di DNA che si leghino alle unità genetiche che codificano per Fn3. Queste «etichette di riconoscimento», utilizzate assieme a una potente tecnica di amplificazione del DNA, la reazione a catena della polimerasi, potrebbero contribuire all'identificazione dei geni nei batteriofagi o in altri vettori.

L'origine del dominio Fn3 è ancora avvolta dal mistero. Fece la sua prima comparsa negli animali? Oppure siamo semplicemente incapaci di identificarne le forme ancestrali con il confronto delle sequenze? In quanto a struttura tridimensionale, Fn3 ha una somiglianza molto stretta con il dominio immunoglobulinico. La cristallografia a raggi X e la risonanza magnetica nucleare hanno permesso di far risalire questo dominio a proteine presenti nei procarioti - tra cui la PapD, una «chaperonina» che aiuta altre proteine ad avvolgersi - e anche a un enzima batterico che digerisce la cellulosa. È interessante che il dominio immunoglobulinico, come è stato definito originariamente, contenga un ponte disolfuro che ne tiene unite le due parti; le forme più primitive, alcune delle quali si conservano ancora nei vertebrati, sono però prive di questo legame. Proprio queste forme primitive sono le più simili al dominio Fn3.

Possiamo prevedere che saranno scoperti altri esempi di moduli catturati e trasferiti. Secondo il nostro conteggio, il dominio Fn3 si trova in circa una proteina animale su 50 (vale a dire 50 su 2500 sequenze animali note, non includendo in questa cifra proteine uguali presenti in specie diverse). Stimiamo che circa 25 moduli, oltre a Fn3, appaiano con grande frequenza nelle proteine animali, mentre più di 100 altri moduli si trovino in collocazioni differenti, ma con una frequenza inferiore a quella del primo gruppo. Tracciare l'albero genealogico e la diaspora di queste unità modulari è una sfida importante, che dovrebbe chiarire ogni aspetto dell'evoluzione di tutti gli esseri viventi.

BIBLIOGRAFIA

DOOLITTLE RUSSELL F., *The Genealogy of Some Recently Evolved Vertebrate Proteins* in «Trends in Biochemical Sciences», 10, n. 6, giugno 1985.

BARON MARTIN, NORMAN DAVID G. e CAMPBELL IAIN D., *Protein Modules* in «Trends in Biochemical Sciences», 16, n. 1, gennaio 1991.

PATTHY LÁZLÓ, *Modular Exchange Principle in Proteins* in «Current Opinion in Structural Biology», 1, n. 3, giugno 1991.

BORK PEER, *Mobile Modules and Motifs* in «Current Opinion in Structural Biology», 2, n. 3, giugno 1992.

BORK PEER e DOOLITTLE RUSSELL F., *Proposed Acquisition of an Animal Protein Domain by Bacteria* in «Proceedings of the National Academy of Sciences», 89, n. 19, 1 ottobre 1992.

Piante a impollinazione acquatica

Ben lungi dall'essere scherzi di natura, come si pensava un tempo, queste piante acquatiche costituiscono uno straordinario esempio di convergenza evolutiva verso strategie di impollinazione più efficienti

di Paul Alan Cox

In una giornata piena di sole dell'ottobre 1787, il botanico Filippo Cavolini attraversava con una barca a remi il Golfo di Napoli, continuando una ricerca che fino ad allora si era dimostrata infruttuosa. Quattro anni prima gli avevano portato esemplari di una specie di alga, *Zostera marina*, completi di frutti. Cavolini aveva intuito subito che si trattava di una pianta particolare: una pianta marina che doveva produrre fiori, a differenza delle alghe a lui note. Egli voleva quindi trovare quei fiori. Cavolini raggiunse il suo obiettivo il giorno in cui scoprì una popolazione di *Zostera* che cresceva nel Golfo. Si emozionò molto nel trovare «una spiga non diversa da quella di una graminacea che galleggiava sull'acqua».

Più tardi, esaminando il polline al microscopio, Cavolini si stupì nel vedere che i granuli non erano piccoli e tondeggianti come quelli delle piante terrestri: «L'ho trovato [il polline di *Zostera*] diverso da quello delle altre piante; infatti è allungato come una piccola anguilla, che, con movimento improvviso e veloce, emette e diffonde in un batter d'occhio il suo sperma». Sulla base di questa osservazione trasse una conclusione sorprendente: *Zostera* doveva riprodursi con impollinazione acquatica.

La scoperta di Cavolini era straordinaria. Dal momento che il polline delle piante terrestri si disintegra nell'acqua, i botanici avevano considerato l'ambiente acquatico nemico dell'impollinazione. Il polline, che trasporta e protegge l'informazione genetica della pianta, è un adattamento indispensabile per le piante terrestri con fiori, le quali utilizzano il vento e gli animali per diffondere i propri gameti sulla terraferma.

La scoperta dell'acqua come vettore d'impollinazione è stata riconsiderata di recente, alla luce delle teorie militari della ricognizione e della fisica della diffusione. I risultati indicano che i sistemi d'impollinazione idrofili non possono essere considerati un semplice scherzo di natura, come si riteneva in passato; le piante a impollinazione acquatica devono invece essere viste come casi convincenti di convergenza evolutiva, dal momento che specie idrofile non imparentate hanno evoluto strategie analoghe che favoriscono una impollinazione efficiente. Nonostante gli specifici adattamenti all'ambiente acquatico, gli antenati di queste specie possono essere ricondotti a piante che hanno ricolonizzato tale ambiente. Per questo motivo le piante idrofile mantengono ancora molti caratteri riproduttivi delle loro controparti terrestri, proprio come le balene e i delfini hanno mantenuto caratteri da mammiferi dopo il loro ritorno all'oceano.

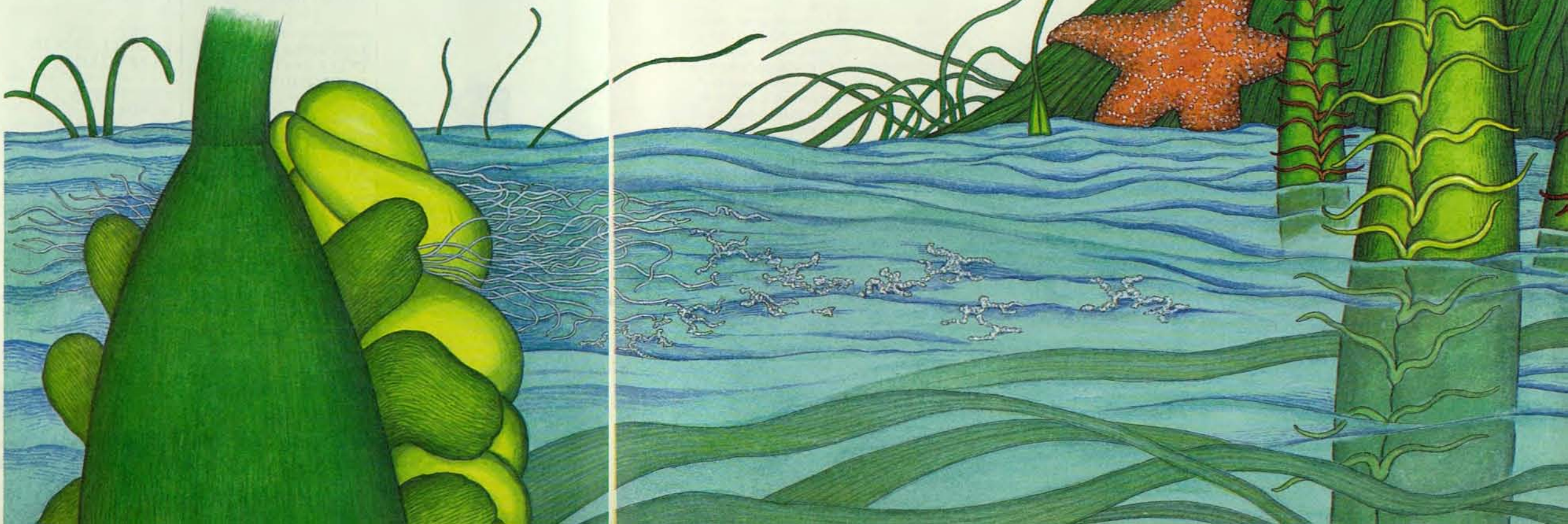
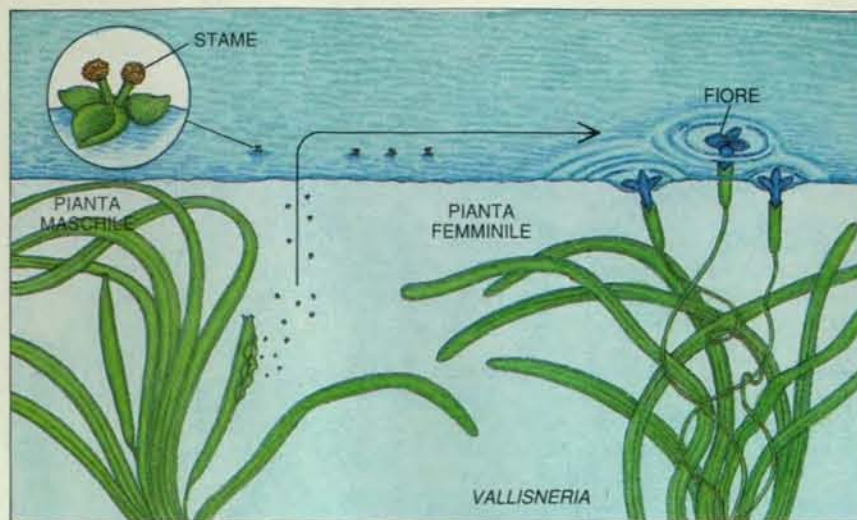
L'impollinazione acquatica non deve essere confusa con la riproduzione di piante prive di fiori come alghe, felci, muschi ed epatiche. Sebbene esse richiedano acqua - gocce di rugiada o spruzzi di un corso d'acqua - non producono polline, ma semplicemente disperdono nell'acqua i loro gameti mobili. L'impollinazione idrofila non è simile neppure all'impollinazione aerea di altre piante acquatiche come le ninfee, che aprono i fiori alla superficie per essere impollinate dagli insetti o dal vento.

Tassonomicamente e geograficamente non rare, le specie a impollinazione acquatica si trovano in 31 generi di 11 famiglie diverse. Sono distribuite in latitudine dalla Svezia settentrionale all'Argentina meridionale e in altitudine da 40 metri sotto il livello del mare a 4800 metri di quota sulle Ande, dove vive *Elodea potamogeton* che ostacola la navigazione nel lago Titicaca. Sono anche adattate ad ambienti molto diversi, dalle foreste pluviali tropicali alle pozze stagionali dei deserti. Le specie di acqua dolce alimentano pesci e uccelli acquatici, mentre quelle marine sostentano tartarughe, lamantini e molluschi.

In che modo le piante idrofile utilizzano come vettore d'impollinazione un mezzo che sembra ben poco adatto per la dispersione del polline? L'acqua tende a far gonfiare e scoppiare la maggior parte dei granuli di polline ed è spesso imprevedibile: le correnti possono essere molto irregolari e le maree possono sommergere all'improvviso le piante o lasciarle esposte all'aria, al di sopra della superficie liquida.

Gli strabilianti adattamenti delle piante acquatiche sono stati notati già da molto tempo. Più di 200 anni fa Erasmus Darwin, nonno di Charles Darwin, descrisse lo straordinario sistema d'impollinazione della pianta d'acqua dolce *Vallisneria*. (La descrizione di Darwin fu così vivida che il reverendo William Paley la citò come prova dell'esistenza di Dio nel suo *Natural Theology*.) Darwin trovò che le piante maschili sommerse liberano fiori che risalgono in superficie e galleggiano. Qui i petali si incurvano all'indietro, sollevando il fiore su un menisco d'acqua che gli consente di spostarsi al minimo soffio di brezza. Due antere sterili fungono da piccole vele, mentre due antere fertili sostengono il polline sferico. I fiori femminili galleggianti, ancorati da un lungo stelo al fondo dello stagno, producono una lieve depressione sulla superficie dell'acqua. Quando il fiore maschile cade in questa depressione, venendo a contatto con il fiore femminile, il polline viene trasferito sullo stimma di quest'ultimo.

Nel 1792 Cavolini sviluppò i propri studi su *Zostera*, dimostrando come nelle piante acquatiche esistesse un altro meccanismo, che consisteva nel liberare il polline direttamente in acqua. Cavolini riferì anche di un'altra monocotiledone marina, *Cymodocea nodosa*, che libera polline nastriforme sulla superficie dell'acqua. Egli scrisse che «il polline bianco e filamentoso... copriva per largo tratto la superficie dell'acqua». Questo polline non scoppiava in quanto era in equi-

Il fiore maschile di *Phyllospadix scouleri* (*a sinistra*) libera polline filamentoso che forma masse galleggianti simili a fiocchi di neve. Le correnti portano queste masse verso i fiori femminili (*a destra*), dove i granuli raggiungono gli stimmi e realizzano così la fecondazione.

L'impollinazione al di sopra della superficie dell'acqua è una strategia riproduttiva cui si affidano alcune specie idrofile. I granuli di polline sono tenuti sempre sopra l'acqua, e vengono trasportati dal fiore maschile galleggiante che si muove verso i fiori femminili per fecondarli. Ne sono esempi *Enhalus*, *Lagarosiphon* e *Vallisneria*.

librio osmotico con l'acqua di mare.

Le osservazioni di Darwin e di Cavolini furono degnate solo di un fugace interesse da parte di altri ricercatori, che continuarono a concentrarsi sull'impollinazione a opera degli animali e del vento. Solo negli ultimi 20 anni i botanici si sono posti domande più ampie, di carattere evolutivo, sull'impollinazione idrofila. Esistono principi generali che si possono applicare a tutti i sistemi d'impollinazione idrofila? Quali aspetti dovrebbero caratterizzare sistemi di impollinazione efficienti su un piano bidimensionale come la superficie dell'acqua? Quale forma e dimensione dovrebbe avere il polline? E i granuli verrebbero dispersi singolarmente o in gruppi?

Queste domande ricadono nell'ambito della teoria della ricognizione sviluppata durante la seconda guerra mondiale dal matematico Bernard O. Koopman e colleghi dell'Antisubmarine Warfare Operations Group della US Navy. Koopman ha dimostrato in maniera semplice, ma elegante che, se un veicolo da ricognizione segue un percorso casuale in due dimensioni, la sua probabilità di incontrare un obiettivo stazionario aumenta rapidamente con la larghezza del percorso. (Matematicamente la relazione è $p = 1 - e^{-wL/A}$, dove p è la probabilità di trovare l'obiettivo, e è la base dei logaritmi naturali, w è la larghezza del cammino percorso dal veicolo da ricognizione, L è la lunghezza del cammino e A è l'area di ricognizione.) L'equazione di Koopman ha un significato militare immediato: un aumento anche piccolo della portata del radar di una nave aumenta notevolmente la probabilità di individuare un obiettivo stazionario durante una ricognizione casuale.

L'equazione di Koopman si applica non solo alle navi, ma anche ai granuli di polline in cerca dello stimma, e consente di definire un limite inferiore per l'efficienza della ricerca da parte dei granuli pollinici quando questi non seguono percorsi del tutto casuali. Se la probabilità di trovare l'obiettivo è bassa (ipotesi ragionevole nell'impollinazione acquatica), allora un piccolo aumento delle dimensioni del veicolo da ricognizione incrementerà sensibilmente la probabilità di raggiungere uno stimma sulla superficie dell'acqua.

Il significato evolutivo dell'equazione di Koopman è chiaro per *Vallisneria*, la pianta studiata da Erasmus Darwin. I fiori maschili galleggianti esplorano un percorso migliaia di volte più largo di quello che può essere seguito da un singolo granulo di polline, aumentando così l'efficienza della ricerca. Un sistema d'impollinazione simile si riscontra in una specie acquatica africana del genere *Lagarosiphon*, studiata da C. D. K. Cook dell'Università di Zurigo. Come in *Vallisneria*, i fiori maschili vengono liberati sott'acqua e si portano in superficie, dove tre stami sterili si drizzano, fungendo da vele in miniatura. I lunghi stami fertili, che si dispongono paralleli alla superficie dell'acqua, portano un polline vischioso all'estremità di filamenti a bastoncino. I fiori maschili si riuniscono in grandi masse galleggianti che vanno a collidere con i fiori femminili, pure galleggianti sull'acqua.

Ma che dire delle piante come quella studiata da Cavolini, *Z. marina*, che liberano il polline direttamente sulla superficie dell'acqua? Anche per esse vi è il vantaggio di un ampio percorso di ricognizione? Per rispondere a queste domande R. Bruce Knox dell'Università di Melbourne e io abbiamo esaminato la biologia dell'impollinazione di diversi generi di piante idrofile del Pacifico meridionale. Una monocotiledone marina da noi studiata è stata *Halodule pinifolia*, che cresce lungo le coste sabbiose delle isole Fiji. Mentre conducevo il lavoro sul campo con P. B. Tomlinson della Harvard University e John A. West dell'Università della California a Berkeley, ho scoperto che intere popolazioni di *Halodule* erano quasi del tutto emerse durante le basse maree primaverili. In quel periodo gli stami bianchi di *Halodule* liberano lentamente lungo polline nastriforme che galleggia sul mare.

I nastri di polline hanno un rivestimento di proteine e carboidrati che li rende vischiosi, così da aderire fra loro, e formano masse galleggianti che assomigliano a fiocchi di neve. Milioni di questi veicoli fiorali da ricognizione vengono trasportati con la marea alle pozze poco profonde dove galleggiano gli stimmi filamentosi delle piante femminili. La collisione dei veicoli da ricognizione con gli stimmi sulla superficie dell'acqua determina l'impollinazione.

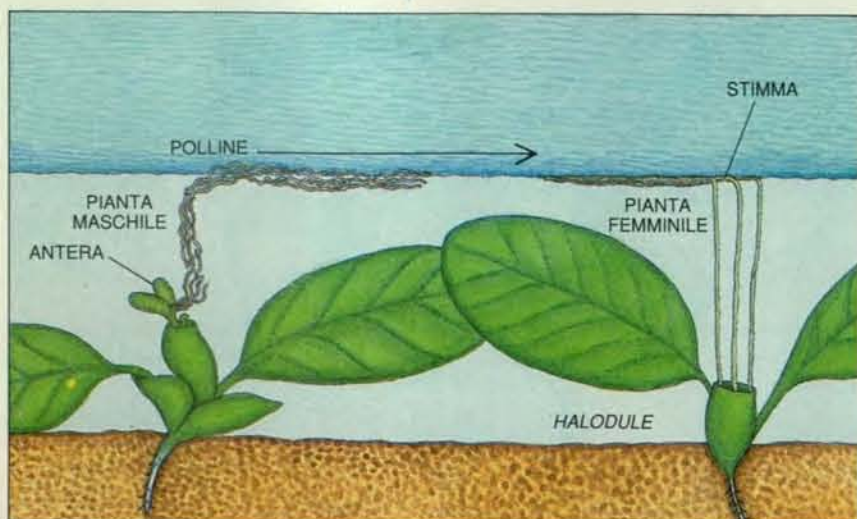
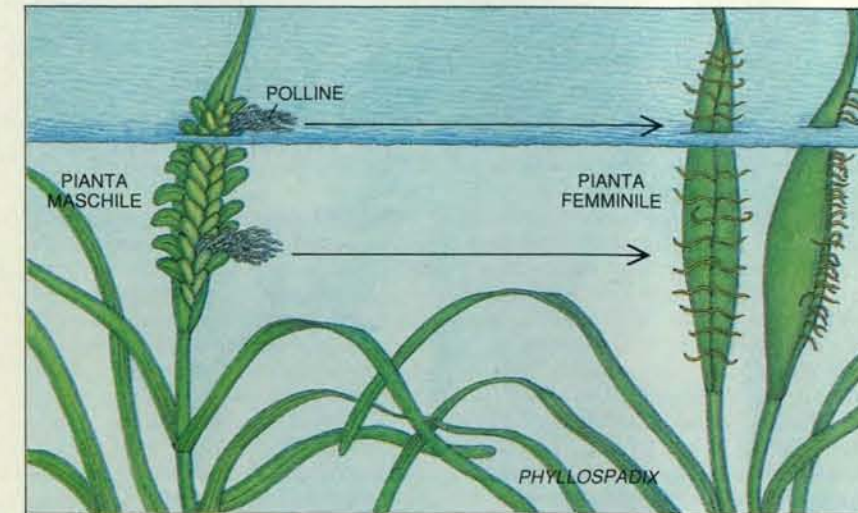
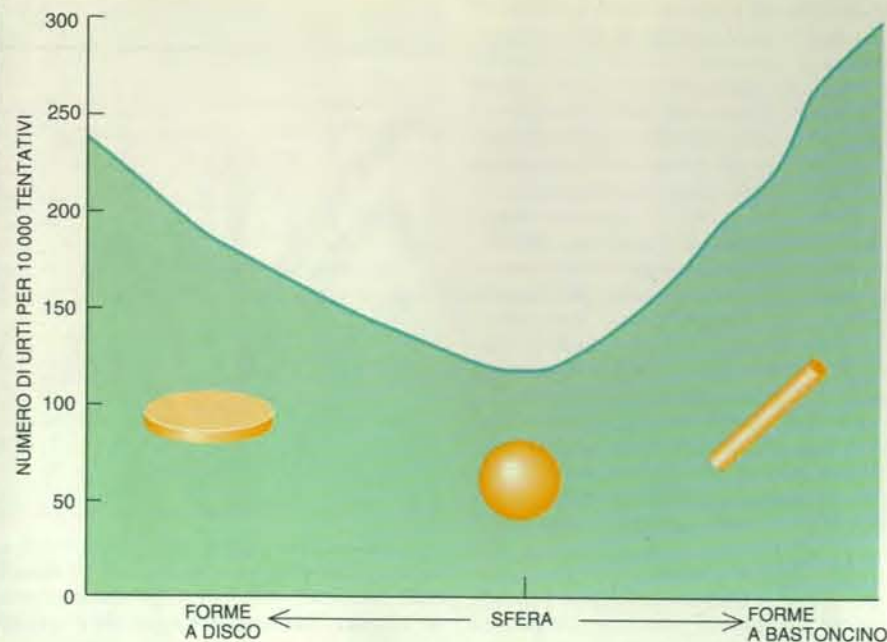
La liberazione di polline sull'acqua è un altro metodo riproduttivo. A seconda delle specie, le masse galleggianti sono formate da granuli ovali o filamentosi, oppure sferici e inseriti in lunghi tubi. Gli stimmi sono filamentosi in piante marine come *Amphibolis* e *Halodule*, e a coppa in specie di acqua dolce come *Ruppia* e *Lepilaena*.

Molti altri generi di monocotiledoni marine si dotano di veicoli da ricognizione che ampliano il percorso di ricerca, benché il meccanismo vari leggermente. In *Amphibolis* dell'Australia e *Thalassodendron* del Kenya i fiori maschili si portano in superficie e liberano polline filamentoso che si riunisce in masse dall'aspetto di graticci. Alcuni generi non imparentati di monocotiledoni marine non hanno polline nastriforme, ma si adattano altrimenti. Nelle Fiji i granuli ovali del polline di *Halophila ovalis* non costituiscono ammassi, ma sono inseriti in vischiosi tubetti mucillaginosi che si uniscono in masse simili a piume galleggianti. In Australia la specie di acqua dolce *Lepilaena cylindrocarpa* libera granuli di polline sferici in grovigli mucillaginosi galleggianti, simili a frittatine, che cadono nelle depressioni create dai fiori femminili.

Perciò appare chiaro che numerose piante idrofile - tra le altre *Vallisneria, Halodule, Halophila* e *Lepilaena* - rappresentano un esempio di convergenza evolutiva: specie non imparentate hanno evoluto meccanismi simili (veicoli da ricognizione galleggianti), che servono ad aumentare l'efficienza dell'impollinazione. Naturalmente, alcune specie idrofile presentano adattamenti del tutto peculiari, meccanismi che potrebbero rappresentare uno scherzo dell'evoluzione. Quale stratega militare non sarebbe incuriosito da *Zannichellia palustris*? Cook, io e i nostri rispettivi studenti You Hao Guo e Rebecca Sperry abbiamo scoperto che *Zannichellia* forma minacciose nubi galleggianti, che lentamente lasciano cadere i granuli di polline, come bombe di profondità, sugli stimmi sottostanti. E quale golfista non sarebbe deliziato dal sistema d'impollinazione di un'altra pianta d'acqua dolce, *Hydrilla verticillata*, studiata anch'essa da Cook? I fiori maschili si portano in superficie e in modo esplosivo catapultano in aria il polline verso i fiori femminili galleggianti. Dato che il polline di *Hydrilla* è distrutto al contatto con l'acqua, contano solo i colpi andati a segno, ma, considerata l'enorme quantità di polline liberato, la probabilità che uno stimma venga colpito è in realtà elevata.

Finora ho trattato dei sistemi d'impollinazione delle piante il cui polline viene trasportato sulla superficie dell'acqua; in questi casi il polline si muove in due dimensioni. Alcune specie hanno sistemi d'impollinazione che operano in tre dimensioni, cioè sotto la superficie (si veda l'articolo *Impollinazione sottomarina* di John Pettitt, Sophie Ducker e Bruce Knox in «Le Scienze» n. 153, maggio 1981).

Le strategie d'impollinazione in tre dimensioni sembrano essere meno comuni di quelle in due dimensioni, forse perché intrinsecamente meno efficienti. In tre dimensioni il polline e gli stimmi sono dispersi in un volume anziché trovarsi su un solo piano, e occorre quindi più polline. Anche i risultati dei moti casuali differiscono notevolmente. Una ricerca casuale è molto più efficiente in due dimensioni, dato che il percorso di ricognizione è ricorrente: prima o poi, qualsiasi stimma che si trovi su un piano verrà incontrato. Non si può dire la stessa cosa per la ricognizione tridimensionale, anche avendo un tempo infinito a disposizione (si veda l'articolo *Moto browniano e teoria del potenziale* di R. Hersch e R. J. Griego in «Le Scienze» n. 10, giugno 1969).

L'impollinazione tridimensionale si è evoluta nelle piante idrofile in cui normalmente l'impollinazione bidimensionale avrebbe scarse opportunità di affermarsi, cioè nelle piante che sono sempre sommerse. Un esempio è costituito da *Thalassia testudinum*, una pianta dei Caraibi che Tomlinson, Thomas Elmqvist dell'Università di Umeå in Svezia e io abbiamo studiato a St. Croix. *Thalassia* libera sott'acqua un polline tondeggiante racchiuso in filamenti di

L'impollinazione subacquea è una terza strategia riproduttiva. In alcune specie il polline, filamentoso o incluso in filamenti mucillaginosi, è trasportato sott'acqua dalle correnti; altre specie producono polline più pesante dell'acqua. Fra gli esempi vi sono *Thalassia, Syringodium* e una specie del genere d'acqua dolce *Lepilaena*.

Una modalità mista di trasporto del polline, adottata da *Phyllospadix* e *Zostera*, unisce l'impollinazione di superficie a quella sottomarina. Il polline liberato in superficie è un po' più filamentoso dell'altro. L'emissione sottomarina si ha nella zona subtidale, in popolazioni che non sono esposte all'aria neppure durante la bassa marea.

La forma ottimale dei granuli per l'impollinazione subacquea è stata studiata con il supercalcolatore. Si sono deformate sfere in forme ellissoidali di uguale volume, allungate come bastoncini o appiattite come dischi. I bastoncini, la forma più efficace nel colpire un bersaglio, sono confrontabili con veri granuli di polline, come quello di *Phyllospadix* (*nella fotografia*).

Per l'impollinazione di superficie la forma ottimale dello stimma è stata determinata deformando sfere in forme ellissoidali sempre più allungate. I granuli di polline collidono più frequentemente con i lunghi stimmi ellittici che con quelli tondeggianti. I risultati sono confrontabili con l'osservazione degli stimmi di piante come *Zostera marina* (*nella fotografia*).

mucillagine. Trascinati sott'acqua dalle onde, i filamenti collidono con i rigidi e setolosi stimmi delle piante femminili. Un meccanismo molto simile d'impollinazione subacquea è caratteristico di una specie di acqua dolce non imparentata con la precedente, *Lepilaena bilocularis*, rinvenuta nel Lago Ellesmere, nell'Isola del Sud della Nuova Zelanda.

Ho voluto verificare se la tecnica messa in atto da queste specie lontane geograficamente e tassonomicamente rappresentasse una strategia tridimensionale efficiente. Insieme con gli studenti di matematica Tyler Jarvis, ora alla Princeton University, e Scott Cromar, ora alla Rutgers University, ho effettuato una serie di esperimenti numerici su un supercalcolatore all'IBM Scientific Computing Center di Palo Alto, in California. Abbiamo deformato ellitticamente in vario grado alcune sfere, mantenendone il volume costante, per produrre una serie di lunghi fili o di dischi appiattiti simili a *frisbee*. Abbiamo poi rilevato con quale frequenza ogni forma urtava bersagli posti casualmente. Dalle simulazioni appariva chiaro che tutte le deformazioni ellittiche erano più efficienti delle sfere nella ricognizione tridimensionale.

Di recente un mio studente, David Smith (ora a Princeton), David A. Olson, allora al Courant Institute of Mathematical Sciences della New York University, e io abbiamo confermato analiticamente questi esperimenti numerici. Le equazioni hanno messo in luce anche qualcosa che gli esperimenti con il supercalcolatore non avevano evidenziato: le forme leggermente appiattite sono le meno efficienti nella ricognizione tridimensionale. Questo però non contraddice la conclusione generale. I veicoli da ricognizione di *T. testudinum* del Mar delle Antille e *L. bilocularis* della Nuova Zelanda - ciascuno costituito da polline attaccato a un filamento mucillaginoso - rappresentano soluzioni di alta efficienza al problema di individuare oggetti sparsi nelle tre dimensioni.

Alcune specie idrofile uniscono strategie bi- e tridimensionali. Per esempio *Enhalus acoroides*, una monocotiledone marina che ho studiato nell'isola indonesiana di Banda, fiorisce solo durante le basse maree primaverili. Come in *Vallisneria*, i bocci fiorali maschili galleggiano sull'acqua, si aprono e vengono catturati tra i tepali arricciati dei fiori femminili; quando la marea si alza, i tepali dei fiori femminili si chiudono, intrappolando i fiori maschili. Tuttavia *Enhalus* disperde anche una parte del polline direttamente sulla superficie dell'acqua.

Anche *Phyllospadix scouleri* disperde polline nastriforme sopra e sotto la superficie. In uno studio eseguito sulla costa settentrionale della California, Tomlinson, il mio studente Kevin Nieznanski e io abbiamo osservato veicoli da ricognizione formati da polline galleggiante, che urtavano gli stimmi femminili. Microfotografie elettroniche di stimmi prelevati molto al di sotto della superficie hanno però rivelato la presenza di polline, indicando che può avvenire anche un'impollinazione sottomarina.

Ancora un altro esempio d'impollinazione mista sottomarina e di superficie si riscontra nella pianta inizialmente studiata da Cavolini: *Z. marina*. A. W. A. M. de Cock dell'Università Cattolica dei Paesi Bassi ha scoperto che, sebbene il polline di *Zostera* liberato sott'acqua in acquario affondi lentamente, il polline disperso durante la bassa marea nelle popolazioni intertidali galleggia in superficie. Roger H. Laushman dell'Oberlin College, Mary H. Ruckelshaus dell'Università di Washington e io, lavorando ai Friday Harbor Marine Laboratories nelle San Juan Islands, al largo dello Stato di Washington, abbiamo confermato l'osservazione di de Cock. Abbiamo infatti trovato masse di polline simili a fiocchi di neve che galleggiavano in superficie, ma anche «mazzi» lineari di polline nastriforme dispersi sott'acqua.

Un metodo di ricognizione efficiente non è l'unico modo per aumentare le probabilità di impollinazione; l'evoluzione può anche selezionare stimmi di facile reperibilità. Sfortunatamente preparare un obiettivo facile da colpire non è mai stato considerato un compito prioritario dagli strateghi militari che hanno sviluppato la teoria della ricognizione; così gran parte dell'analisi di questo modello è toccata ai botanici.

Dal punto di vista dello stimma, esistono almeno due possibili soluzioni per aumentare la probabilità di essere raggiunto da un granulo di polline o da una massa pollinica: ampliare l'area utile del bersaglio oppure modificare la forma dello stimma per incrementare la probabilità della collisione.

Le strutture fiorali di molte piante a impollinazione acquatica sono un esempio della prima soluzione. Alcuni fiori, come per esempio quelli di *Vallisneria* e di *Enhalus*, creano depressioni sulla superficie dell'acqua, formando un'area bersaglio molto più ampia di uno stimma o anche di un intero fiore femminile. Un ulteriore esempio è fornito da *Lepilaena cylindrocarpa* in Australia. Usher Posluszny, dell'Università di Guelph, ha attirato la mia attenzione su una piccola brattea che tiene uniti, prima della fecondazione, tre stimmi indipendenti. Così ravvicinati, questi tre stimmi, ciascuno dei quali assomiglia a un guantone da baseball, provocano una depressione sulla superficie dell'acqua.

Anche il movimento oscillatorio di alcune parti della pianta può ampliare l'area bersaglio. In acque che si muovono lentamente lo stelo allungato di *Ruppia marina* funziona come un tergicristallo: fa sì che i fiori ondeggino delicatamente avanti e indietro e raccolgano le masse galleggianti di polline. Inoltre le strutture fiorali possono avere effetti localizzati. Joseph D. Ackerman, allora alla Cornell University, aveva osservato come le infiorescenze sommerse di *Zostera* potessero modificare il flusso dell'acqua, trovando che, in condizioni di laboratorio, il cambiamento di flusso poteva concentrare il polline vicino agli stimmi.

Lo studio della seconda possibile strategia di potenziamento della visibilità dell'obiettivo - cioè un cambiamento di forma dell'obiettivo stesso - è iniziato solo negli ultimi anni. In una serie di simulazioni al supercalcolatore, James A. Sethian di Berkeley e io abbiamo trovato che in due dimensioni gli stimmi filamentosi costituiscono un bersaglio molto più facile di quelli tondeggianti. Gli stimmi lunghissimi e filiformi delle monocotiledoni marine dei generi *Halodule*, *Halophila* e *Thalassodendron* sembrano confermare la previsione.

Le osservazioni sul campo contraddicono anche le ipotesi formulate in precedenza da alcuni ricercatori, secondo cui le piante idrofile fiorirebbero e si impollinerebbero solo raramente. Ho trovato invece che la fioritura e l'impollinazione incrociata (il trasferimento di polline da una pianta a un'altra) sono eventi comuni, anche se talora effimeri. La frequenza della fioritura - e di conseguenza dell'impollinazione incrociata - può influenzare direttamente la diversità genetica delle popolazioni. Utilizzando la tecnica di elettroforesi sul gel d'amido, Laushman ha scoperto che nelle popolazioni di *Zostera* l'impollinazione incrociata è relativamente frequente, in particolare nella zona intertidale dove l'impollinazione avviene in superficie.

Altri ricercatori hanno però trovato ben poca diversità. In base ai suoi studi sulla specie di acqua dolce *Ceratophyllum demersum*, Donald H. Les della Università del Wisconsin a Milwaukee pensa che le piante idrofile possano essere in gran parte geneticamente uniformi; l'uniformità avrebbe origine dall'autoimpollinazione. Per esempio, C. Thomas Philbrick del Rancho Santa Ana Botanic Garden di Claremont, in California, ha scoperto che in una specie sommersa di *Potamogeton* minuscole bolle d'aria possono trasportare il polline dall'antera allo stimma dello stesso fiore. L'uniformità genetica può risultare anche da fioriture infrequenti e dall'ampiezza dei cloni in una popolazione.

Sebbene la dinamica dell'impollinazione acquatica non sia valsa a fornire dettagli sulla genetica delle popolazioni idrofile, essa fornisce indizi per spiegare come mai queste piante abbiano seguito una via evolutiva diversa da quella delle loro controparti terrestri. Il polline nastriforme è un carattere distintivo delle piante idrofile. Per studiare la comparsa di questo polline C. J. Humphries del London Natural History Museum e io abbiamo tentato di ricostruire l'albero genealogico della famiglia di monocotiledoni marine *Cymodoceaceae*. Abbiamo così scoperto che questa famiglia, che comprende *Thalassodendron*, *Amphibolis*, *Syringodium*, *Halodule* e *Cymodocea*, ha un antenato comune con *Posidonia* (monocotiledone marina che vive nel Mediterraneo e nell'Australia occidentale) e con le *Zosteraceae*, famiglia che comprende *Zostera*, *Phyllospadix* e il genere australiano *Heterozostera*. La nostra ricerca fa pensare che questo antenato comune avesse polline nastriforme. Perciò, probabilmente, il polline nastriforme si è evoluto una sola volta, e ciò spiega perché altri generi di monocotiledoni marine non imparentate, come *Halophila*, *Thalassia* ed *Enhalus*, conservino ancora polline tondeggiante simile a quello delle piante terrestri.

Un'altra caratteristica singolare delle piante idrofile riguarda il dioicismo, ossia la separazione di fiori maschili e femminili su piante distinte. Più della metà delle specie idrofile è dioica, mentre quelle terrestri lo sono soltanto per il 3-4 per cento. Il dioicismo nelle piante idrofile potrebbe essersi evoluto come

Solo nel corso delle basse maree primaverili molte piante marine, come questa popolazione di *Phyllospadix* lungo la costa di Monterey, in California (*a sinistra*), hanno l'opportunità di esporre il polline o i fiori alla superficie dell'acqua. Il polline di numerose specie idrofile tende ad aggregarsi in masse galleggianti, come avviene nel caso di *Amphibolis* (*a destra*).

conseguenza dell'impossibilità fisica di coesistenza in un stesso fiore dei meccanismi di dispersione del polline operata dall'acqua e di cattura del medesimo. Una morfologia fiorale mista farebbe probabilmente sì che il fiore catturasse il proprio polline prima di liberarlo, annullando in questo modo la capacità della pianta di effettuare un'impollinazione incrociata.

Un'altra possibile spiegazione dell'elevata percentuale di dioicismo osservata è che tale processo impedisce l'inimpollinazione. Dal momento che molte popolazioni idrofile sono costituite da grossi cloni, si può verificare inimpollinazione quando le parti maschili e femminili del fiore si trovano sulla stessa pianta. I meccanismi alternativi di esoimpollinazione che sono tipici delle piante terrestri probabilmente non funzionerebbero nelle piante idrofile. Per esempio, la maturazione dei fiori maschili e femminili in tempi diversi, fenomeno caratteristico delle piante terrestri, sarebbe probabilmente difficile da sincronizzare in un grosso clone di monocotiledoni marine.

Questi caratteri insoliti rendono le piante idrofile non solo affascinanti, ma anche estremamente utili per lo studio di numerosi problemi riguardanti l'ecologia evolutiva e la demografia delle piante. La durata della vita del polline idrofilo è facile da stabilire, dal momento che i granuli pollinici di molte piante idrofile mostrano correnti citoplasmatiche. Questo carattere, facilmente osservabile al microscopio, consiste semplicemente nel movimento del citoplasma all'interno del granulo di polline. Ho trovato che il polline di *Zostera*, per esempio, ha una vita sorprendentemente breve: muore quasi totalmente entro otto ore dalla liberazione. Al contrario, determinare la durata della vita del polline terrestre è molto più difficile: di solito è indispensabile colorarlo per determinarne la vitalità.

Inoltre la sorte dei singoli granuli nelle specie idrofile a impollinazione di superficie è facilmente osservabile. È sufficiente esaminare, immagine per immagine, registrazioni video della superficie dell'acqua e seguire il percorso dei granuli. Questa tecnica è pressoché inattuabile con le piante il cui polline è trasportato dal vento o dagli animali.

Le specie idrofile possono offrire indizi nell'interpretazione dei modelli di colonizzazione geografica e nella genetica di popolazione delle piante, dal momento che le popolazioni di piante idrofile d'acqua dolce sono in sostanza riproduttivamente isolate. Il polline trasportato dall'acqua non può spostarsi da una pozza all'altra. Il flusso genico tra popolazioni deve perciò verificarsi soprattutto per trasporto di frutti o di piccole parti vegetative che si staccano. Lo sviluppo clonale vigoroso della maggior parte delle piante a impollinazione acquatica le rende ideali per esperimenti di ecologia: si può esporre lo stesso genotipo a numerosi ambienti diversi.

In rapporto all'assai diffuso interesse per l'impollinazione operata dal vento e dagli animali, ben poco lavoro è stato compiuto sull'impollinazione idrofila da quando Cavolini, per primo, scoprì fiori di *Zostera* nel Golfo di Napoli. Anche se le piante a impollinazione acquatica non producono fiori grandi e vistosi come molte piante terrestri, personalmente le trovo davvero belle. Queste piante offrono eccellenti opportunità ai botanici sia dilettanti sia professionisti per condurre osservazioni originali sulla storia naturale e per studiare più da vicino la convergenza evolutiva nell'ambiente acquatico.

BIBLIOGRAFIA

KOOPMAN B. O., *Search and Screening*, Pergamon Press, 1980.

COX P. A., *Hydrophilous Pollination* in «Annual Review of Ecology and Systematics», 19, pp. 261-280, 1988.

COX P. A. e KNOX R. B., *Two-Dimensional Pollination in Hydrophilous Plants: Convergent Evolution in the Genera Halodule (Cymodoceaceae), Halophila (Hydrocharitaceae), Ruppia (Ruppiaceae), and Lepilaena (Zannichelliaceae)* in «American Journal of Botany», 76, n. 2, febbraio 1989.

COX PAUL A., LAUSHMAN ROGER H. e RUCKELSHAUS MARY H., *Surface and Submarine Pollination in the Seagrass* Zostera marina L. in «Botanical Journal of the Linnaean Society», 109, n. 2, giugno 1991.

COX P. A., CROMAR S. e JARVIS T., *Underwater Pollination, Three-Dimensional Search, and Pollen Morphology: Predictions from a Supercomputer Analysis* in *Pollen and Spores: Patterns of Diversification*, a cura di S. Blackmore e S. H. Barnes, Oxford University Press, 1992.

Il recupero della *Vasa*

Può dirsi finalmente concluso il lunghissimo e complesso lavoro di restauro che ci ha restituito l'ammiraglia della marina da guerra svedese, rimasta per ben tre secoli sui fondali del porto di Stoccolma

di Lars-Åke Kvarning

Negli anni immediatamente successivi al 1620, la Svezia di Gustavo II Adolfo era avviata a diventare una delle maggiori potenze d'Europa. Controllava la Finlandia, l'Estonia e la Livonia e aveva appena conquistato la piccola parte della Russia che si affacciava sul Golfo di Finlandia. Il Mar Baltico era così diventato un lago svedese, ma il consolidamento di questo predominio richiedeva la presenza di una flotta potente. Pertanto, nel gennaio 1625, il sovrano ordinò la costruzione di quattro navi, la più imponente delle quali doveva essere la *Vasa*; così chiamata in onore della famiglia reale svedese, era una delle più grandi navi da guerra dell'epoca.

Il 10 agosto 1628, in una bella giornata di sole, una grande folla si radunò nel porto di Stoccolma per assistere alla partenza del nuovo vascello, diretto verso una base navale su un'isola poco lontana dalla città. I membri dell'equipaggio spiegarono quattro delle 10 vele della *Vasa*: quelle di trinchetto, di parrocchetto, di gabbia e di mezzana. I cannoni spararono a salve e la folla radunata sui moli e a riva esultò nel vedere l'imponente veliero che cominciava a muoversi lentamente nella brezza leggera, risplendente di colori e dorature.

All'improvviso le acclamazioni si trasformarono in grida di orrore perché la nave, ancora in porto, venne sbilanciata da una folata di vento; si piegò verso babordo e si rialzò un poco, per inclinarsi subito dopo, con l'acqua che irrompeva attraverso le cannoniere aperte. Poi la *Vasa* affondò, «con vele, pennoni e tutto» come si legge nella lettera in cui il Consiglio del regno informava il re del disastro. Secondo l'ambasciatore danese, Erik Krabbe, «si dice che siano annegate più di cinquanta persone, fra cui anche le mogli e i figli di alcuni membri dell'equipaggio, che volevano seguire i congiunti nell'arcipelago».

A oltre tre secoli di distanza, il relitto della *Vasa* è stato recuperato e restaurato, un'impresa questa che fa della nave un monumento tanto ai suoi costruttori quanto ai suoi salvatori. La resurrezione della *Vasa* è importante per l'interesse intrinseco dell'operazione, ma anche per le informazioni che può darci sui suoi costruttori. Una nave non è solo una società in miniatura; è anche il prodotto e il riflesso di una società più vasta.

La perdita di una delle più grandi navi da guerra d'Europa, nel viaggio inaugurale e in acque sicure, era una tale catastrofe da richiedere un capro espiatorio; pertanto l'interrogatorio del capitano, Söfring Hansson, e degli altri ufficiali cominciò subito dopo il loro salvataggio. Il costruttore, Henrik Hybertson, era morto l'anno precedente, e la responsabilità di completare l'opera era stata assunta dalla vedova e dal fratello di costui, in collaborazione con l'assistente, Hein Jacobson; tutti dovettero sottoporsi a loro volta a interrogatorio.

Vi era il dubbio che i cannoni non fossero stati bloccati adeguatamente; l'inchiesta non poté stabilirlo, ma nell'estate 1961 gli archeologi subacquei trovarono che effettivamente gli affusti dei cannoni erano ai loro posti, con i resti delle funi di fissaggio ancora avvolti intorno agli assali. Anche per quanto riguardava le pietre della zavorra, un testimone assicurò che lo spazio appositamente riservato nella stiva era stato riempito.

Tuttavia è possibile leggere fra le righe alcuni segni premonitori. Il nocchiero Göran Mattson riferì che, mentre la nave veniva armata, l'ammiraglio Klas Fleming e il capitano della *Vasa* ne avevano sperimentato la stabilità facendo correre 30 uomini avanti e indietro sul ponte di coperta. Al terzo passaggio avevano dovuto ordinare loro di fermarsi per impedire che la *Vasa* si rovesciasse sul posto. C'era chiaramente qualcosa che non andava nel progetto, ma non fu fatto nulla per rimediarvi.

Dai documenti si può dedurre che, secondo la maggior parte degli osservatori, il compartimento per la zavorra non era abbastanza grande per equilibrare i due ponti di batteria voluti dal re. In effetti, i costruttori sostennero di aver realizzato il vascello secondo i termini del contratto e in accordo con gli ordini ricevuti direttamente dal sovrano.

Gustavo Adolfo voleva una marina forte per proteggere le linee di rifornimento verso il teatro di guerra sulla sponda opposta del Baltico. Intendeva anche bloccare i ricchi porti polacchi, in modo che i suoi agenti potessero esigere tributi dalle navi mercantili qui dirette. Sembra che le sue ambizioni siano cresciute dopo che egli ebbe dato l'ordine di costruire la *Vasa*, e che la nave sia diventata più grande di quanto fosse previsto inizialmente.

Questo intervento reale fu forse la vera causa del disastro. I costruttori navali dell'epoca non disegnavano progetti dettagliati e non avevano modo di calcolare la stabilità delle loro creazioni: la loro unica guida era il metodo per tentativi ed errori. Quando ingrandirono la *Vasa* aggiungendo un secondo ponte di batteria, essi lasciarono nella stiva solo lo spazio sufficiente per circa 121 tonnellate di zavorra, ossia meno della metà di quella necessaria. Se però avessero cercato di zavorrare maggiormente la nave, il ponte di batteria inferiore si sarebbe abbassato fino a trovarsi pericolosamente vicino alla linea di galleggiamento.

Molti furono testimoni dell'interesse personale del re per la costruzione del

Lo splendore della *Vasa* è evidente nella magnifica decorazione della poppa, ora restaurata, in cui spiccano simboli reali quali la corona, lo stemma e i grifoni.

Come avvenne il salvataggio della *Vasa*

Sommozzatori muniti di potenti idranti scavarono sei gallerie nei sedimenti sottostanti il relitto (*riquadro*). Fecero poi passare cavi d'acciaio nelle gallerie e ne fissarono le estremità a file parallele di pontoni semisommersi, formando una solida ingabbiatura. Sollevando i pontoni lo scafo venne alzato di circa 2,5 metri dal fondo e trainato verso riva. Quando esso toccò di nuovo il fondo, si abbassarono i pontoni, si tesero i cavi e si ripeté il processo. Occorsero 18 fasi per raggiungere acque abbastanza basse; qui, dopo due anni di preparazione, martinetti montati su pontoni portarono infine la nave in superficie.

vascello. Il maestro d'ascia, interrogato su chi o cosa fosse responsabile del disastro, sospirò: «Dio solo lo sa...». E dato che Dio e il re erano coinvolti fino a questo punto, vi era una sola via d'uscita: il caso fu archiviato.

Una nave di queste dimensioni, equipaggiata con cannoni di bronzo che valevano una fortuna, non poteva non diventare oggetto di tentativi di recupero. Appena tre giorni dopo il naufragio, il Consiglio del regno autorizzò l'inglese Ian Bulmer a mettersi all'opera, ma i risultati furono scarsi: sebbene i principi di queste operazioni fossero ben noti, la *Vasa* era troppo pesante per essere sollevata con i mezzi disponibili all'epoca.

Lo svedese Albrect von Treileben entrò in scena nel 1663, cinque anni dopo aver ideato la campana subacquea che rendeva i siti sommersi molto più accessibili. Così equipaggiato, intraprese con un gruppo di aiutanti il recupero dei cannoni e di altri oggetti amovibili di valore. L'operatore indossava abiti di cuoio per proteggersi dal freddo e saliva su una piattaforma appesa sotto la campana, con la testa e il torace che sporgevano nella bolla d'aria. Poi la campana veniva abbassata fino al relitto.

Francesco Negri, sacerdote ed esploratore ravennate, lasciò ai posteri nel suo *Viaggio settentrionale* (pubblicato postumo nel 1700) questa testimonianza di prima mano delle operazioni di von Treileben:

«Dato poi il segno, e tirato ad alto, portò di sopra afferrata coll'uncino una grossa tavola di rovere con grosse e pesanti lamine di ferro, dopo d'essersi trattenuto sott'acqua un buon quarto d'ora. Gli dimandai per interprete, se avesse potuto starci più: mi rispose che fino a mezz'ora, non più, per riscaldar poi troppo quell'aria ivi chiusa con l'alito, come credo (ma mi dimenticai di domandargliela), se pur non fu per patir troppo freddo alle gambe e coscie attuffate nell'acqua, perché effettivamente tremava, benché nativo del paese e robusto e usato agli strapazzi del corpo. Era allora circa la fine di ottobre di quest'anno prossimo passato 1663...»

In immersione, l'operatore non poteva quasi muoversi e la visibilità di cui godeva era quasi nulla. Tuttavia il gruppo di von Treileben riuscì a riportare a galla la maggior parte dei 64 cannoni della *Vasa*, compresi 48 pezzi per proiettili da 24 libbre, del peso di oltre una tonnellata ciascuno. Allora c'era poco interesse per il resto del relitto, che venne quindi abbandonato in mare; così terminò quella che fu certamente la più straordinaria operazione subacquea condotta in epoca premoderna.

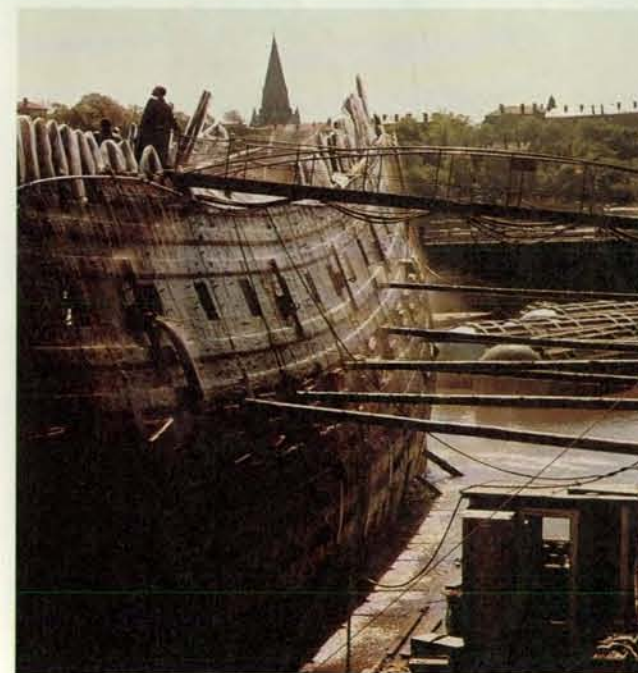
Gli uomini di von Treileben furono fortunati oltre che coraggiosi. La ridotta quantità d'aria disponibile all'interno della campana limitava il tempo di immersione in misura sufficiente a impedire che quantità pericolose di azoto si dissolvessero nel sangue a pressione elevata. In effetti se ciò fosse accaduto, un ritorno troppo rapido alla superficie avrebbe causato un'embolia gassosa probabilmente mortale.

La *Vasa* fu più o meno dimenticata fino al 1956 quando un ricercatore indipendente, Anders Franzén, localizzò il relitto sul fondo del porto di Stoccolma. Il successo di Franzén non fu dovuto al caso, ma a un intento preciso. Egli notò che il Mar Baltico non ha una salinità sufficiente a permettere la vita della teredine (*Teredo navalis*), un mollusco che scava gallerie nelle costruzioni lignee sommerse; nelle sue acque dovrebbero quindi trovarsi relitti straordinariamente ben conservati di navi in legno. Da allora si è effettivamente stabilito che le ricchezze del Baltico sono senza eguali, perché nessun altro mare di bassa salinità è mai stato sede di un traffico navale così intenso.

Franzén si concentrò sulle navi da guerra del XVI e XVII secolo. Consultando gli archivi trovò indicazioni della posizione di 12 relitti che riteneva particolarmente interessanti, e quindi decise di mettersi in cerca della *Vasa*. Un giorno, scandagliando il porto di Stoccolma, si imbatté in un oggetto grande e immobile; un'immersione confermò che si trattava di un enorme scafo di quercia che poggiava ritto sul fondo. Attorno alla *Vasa* e dentro di essa erano sparsi gli scheletri di 25 persone; alcuni vennero scoperti nello scavo del relitto, ma la maggior parte fu trovata dai sommozzatori che esplorarono il fondo dopo che lo scafo fu sollevato.

Il sogno di Franzén di recuperare la nave fu deriso da più parti, ma egli riuscì a procurarsi due sostenitori importanti. La Marina svedese accettò di organizzare l'addestramento dei suoi sommozzatori tenendo presenti le necessità del progetto, e la Broströms Shipping Company, tramite una sua sussidiaria, la Neptun Salvaging Company, si fece carico dell'operazione e di tutte le spese.

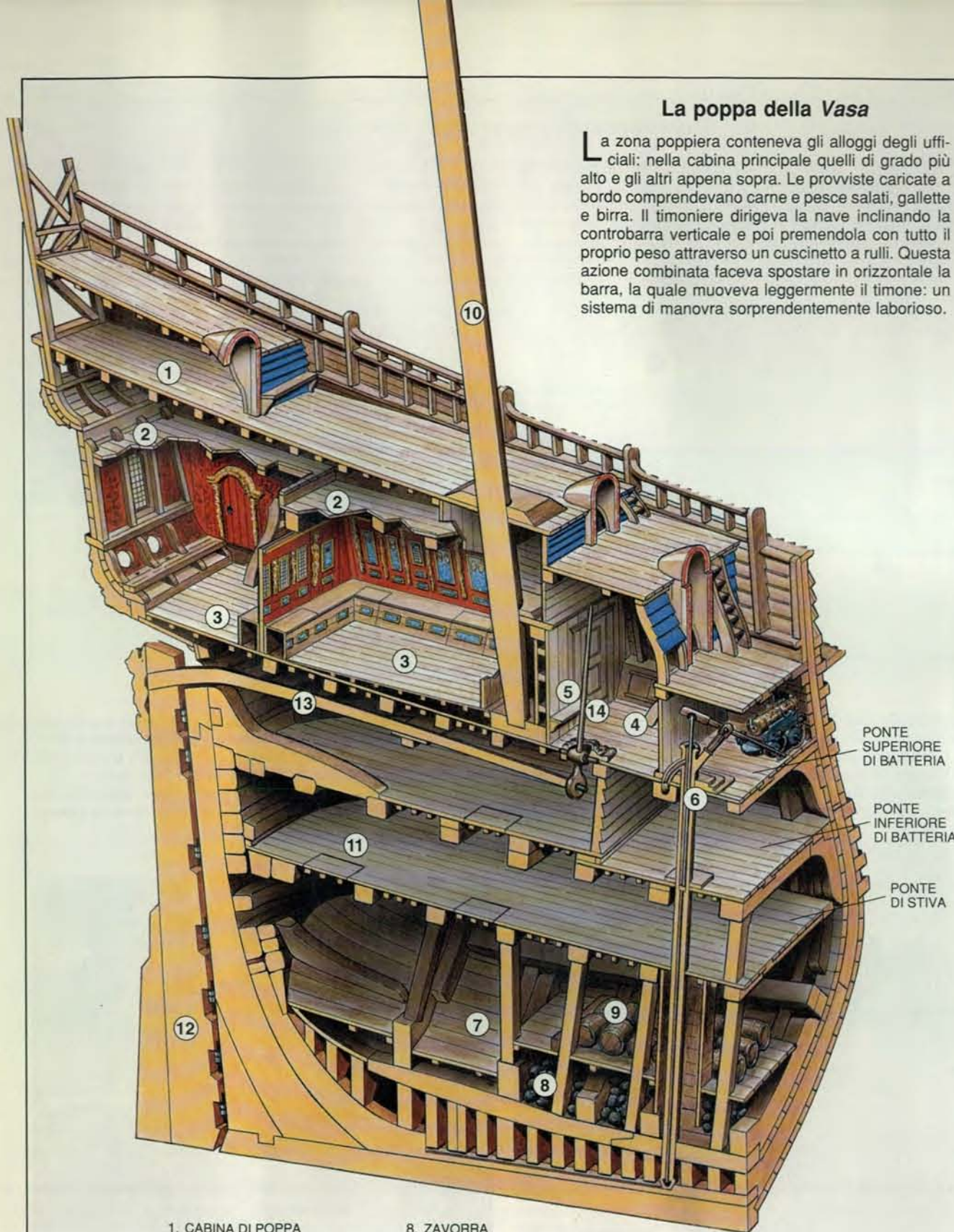
Non mancarono i suggerimenti fantasiosi: fu proposto di riempire la *Vasa* di palline da ping pong per sollevarla in superficie, oppure di congelare l'acqua tutto intorno alla nave in modo che il risultante blocco di ghiaccio salisse a galla, dove avrebbe potuto essere trainato verso la riva per essere fatto fondere.

Alla fine, però, venne prescelta una tecnica tradizionale. Si fecero passare

Lo scafo della *Vasa*, al momento del recupero, era ancora tenuto insieme dalle caviglie in legno originarie, mentre i chiodi e i perni in ferro si erano decomposti in ruggine da molti anni.

Per ben 18 anni lo scafo della *Vasa* venne impregnato con polietilenglicol spruzzato in getti per impedire al legno di deformarsi. Gli oggetti più piccoli furono trattati per immersione.

La poppa della *Vasa*

La zona poppiera conteneva gli alloggi degli ufficiali: nella cabina principale quelli di grado più alto e gli altri appena sopra. Le provviste caricate a bordo comprendevano carne e pesce salati, gallette e birra. Il timoniere dirigeva la nave inclinando la controbarra verticale e poi premendola con tutto il proprio peso attraverso un cuscinetto a rulli. Questa azione combinata faceva spostare in orizzontale la barra, la quale muoveva leggermente il timone: un sistema di manovra sorprendentemente laborioso.

PONTE SUPERIORE DI BATTERIA
PONTE INFERIORE DI BATTERIA
PONTE DI STIVA

1. CABINA DI POPPA
2. ALLOGGI DEGLI UFFICIALI
3. CABINA PRINCIPALE
4. CABINA DI MANOVRA
5. POSTO DEL TIMONIERE
6. POMPA DI SENTINA
7. SANTABARBARA
8. ZAVORRA
9. CAMBUSA
10. ALBERO DI MEZZANA
11. POSTO DEL BARBIERE-CHIRURGO
12. TIMONE
13. BARRA DEL TIMONE
14. CONTROBARRA VERTICALE

cavi sotto lo scafo della *Vasa*, attraverso gallerie scavate nei sedimenti del fondo tramite idranti a doppio getto: il getto anteriore erodeva il materiale del fondo mentre i più deboli getti posteriori servivano a compensare il rinculo e ad allontanare le particelle in sospensione. I sommozzatori dovevano lavorare all'interno delle gallerie, con quattro metri di fanghiglia, una nave vecchia di 333 anni e 32 metri d'acqua direttamente sopra la testa; il pensiero più terrificante, mentre si affaccendavano nell'oscurità totale, era che il fondo dello scafo cedesse sotto il peso della zavorra, intrappolandoli nelle gallerie.

Fu un momento emozionante quando la *Vasa* si trovò racchiusa nella sua ingabbiatura, pronta per lo «strattone» che l'avrebbe liberata dalla morsa dei sedimenti. Due gruppi di pontoni a cui erano fissate le estremità dei cavi vennero svuotati e, sollevandosi, trascinarono i cavi fino a che questi furono tesi come corde di violino. Alla fine il sedimento cedette e la *Vasa* rimase sospesa nella sua ingabbiatura. Questa operazione impose una sollecitazione terribile allo scafo, che si sarebbe gravemente danneggiato se le antiche travi di quercia non avessero conservato ancora una solidità straordinaria. A questo punto l'operazione continuò senza difficoltà.

Dato che i pontoni potevano sollevare la nave solo di 2,5 metri per volta, furono necessarie 18 ripetizioni del procedimento. La *Vasa* era sollevata un poco e trainata in acque più basse, fino a che veniva di nuovo ad adagiarsi sul fondo. Un fatto curioso è che le impronte dello scafo, lasciate sul fondo a ogni fase, sono chiaramente visibili ancora oggi, simili alle orme di un gigante, nei rilevamenti con il sonar.

Alla fine la nave raggiunse un punto in cui il sollevamento finale poteva essere eseguito facilmente. Grandi martinetti montati sui pontoni riportarono a galla il relitto, e pompe che erano state installate nello scafo cominciarono a svuotarlo. Così assistita, la nave galleggiò fino al bacino di carenaggio, dove la attendeva un pontone in cemento. Qui, nel 1961, la *Vasa* rivelò per la prima volta dopo 333 anni la sua mole imponente a un pubblico sbalordito.

Per proteggere il legno, intriso d'acqua e quindi molto delicato, il vascello venne costantemente spruzzato con acqua, in attesa che fosse possibile iniziare il restauro definitivo in un sito appropriato. Era anche necessario asportare la grande quantità di fanghiglia nera che si era accumulata nello scafo, ricca di reperti archeologici e, forse, di batteri. Per prudenza, gli archeologi che dovevano scavare l'interno del vascello vennero vaccinati contro il tetano, il tifo e altre possibili infezioni. Così protetti, essi scoprirono e catalogarono circa 25 000 oggetti, quasi metà dei quali appartenenti alla struttura della nave.

Tutti i reperti dovevano tornare al proprio posto durante l'opera di ricostruzione, il cui obiettivo era quello di riportare la *Vasa* il più vicino possibile alla configurazione originale. Nell'autunno 1961 un riparo in alluminio venne costruito intorno alla nave, che poggiava ancora sul pontone. Questo fu poi trainato in un museo temporaneo, dove i lavori potevano procedere sotto gli occhi del pubblico. Il problema maggiore era quello di sostituire con qualcosa di più durevole l'acqua che intrideva il legno, senza che questo potesse fessurarsi, contrarsi o deformarsi. Era necessario conservare in questo modo non solo 1100 tonnellate di componenti strutturali, ma anche 700 sculture e ornamenti intagliati, oltre a molte migliaia di oggetti in tessuto, cuoio e metallo.

Nessuno aveva mai tentato di trattare una massa così enorme di materiale. Per fortuna Bertil Centervall e Rolf Morén, rispettivamente restauratore e ingegnere, avevano inventato un procedimento per conservare il legno fresco impregnandolo di polietilenglicol (PEG). Dato che il polimero sostituiva l'acqua nel legno fresco, si sperava che potesse fare la stessa cosa con il legno antico intriso d'acqua; di fatto si vide che il metodo funzionava. A questo scopo si preparava una soluzione acquosa di PEG, con aggiunta di borace e acido borico come fungicidi, e la si applicava al legno; il PEG funziona irrigidendo le cellule legnose e impedendo al materiale di cambiare forma o fessurarsi. Ulteriore protezione era fornita da sistemi di climatizzazione che garantivano temperatura costante e un livello ottimale di umidità.

Gli oggetti in legno di dimensioni relativamente piccole venivano immersi in grandi vasche dove la concentrazione e la temperatura della soluzione potevano essere aumentate gradualmente. Questa forma di trattamento molto efficiente richiese 18 mesi per la quercia e circa un anno per i legni dolci. Lo scafo, tuttavia, non poteva essere trattato in questo modo, e si dovette spruzzarlo di soluzione all'interno e all'esterno, in un tedioso processo di saturazione che iniziò nel 1961, tre anni prima che entrassi a far parte del gruppo di restauratori, e terminò solo nel 1979. A questo punto lo scafo venne fatto asciugare lentamente.

Un interessante problema di conservazione venne posto dalle sei vele rimaste (su 10 originarie), scoperte in un gavone (stipo) nel ponte di stiva, il più basso dei quattro ponti che si estendono con continuità da prua a poppa. Il tessuto era così deteriorato che rischiava di cadere in polvere al minimo tocco; era possibile, tuttavia, manipolarlo fino a che lo si teneva in acqua. Vennero dunque costruite grandi vasche nelle quali i resti delle vele potevano essere distesi, puliti e asciugati in alcool e xilene. A causa della loro fragilità, vennero montati su un supporto di fibra di vetro per mezzo di una soluzione plastica appositamente formulata. La soluzione ha lo stesso indice di rifrazione della fibra di vetro e quindi la rende invisibile.

Le parti in ferro puddellato si erano da tempo decomposte in ruggine, con l'eccezione della coppia agugliotto/femminella più bassa (con questi termini si intendono rispettivamente i cardini e le

La costruzione di questo modello della *Vasa*, in scala 1:10, impegnò quattro artigiani per ben quattro anni. Oggi è esposto accanto alla nave restaurata nel Museo Vasa.

I piccoli oggetti ci consentono di farci un'idea molto chiara della vita a bordo della *Vasa*. La vivace decorazione del vascello è testimoniata da questo stemma restaurato, dove il covone rappresenta la parola «Vasa», il nome della famiglia reale. Altri oggetti recuperati includono (*in senso orario da sinistra*) vestiario, corredi per la mensa, articoli per la tavola degli ufficiali, strumenti da farmacista, arnesi da velaio, munizioni, giochi, monete e oggetti per la toeletta.

ghiere che fissavano la pala del timone al dritto di poppa), che era stata coperta dai sedimenti del fondo e si era conservata abbastanza bene. Nel 1992 trovammo anche un perno nascosto in ottime condizioni mentre ci preparavamo a risistemare gli alberi nella scassa.

Le palle di cannone in ghisa avevano resistito meglio all'ossidazione a causa del loro basso contenuto di carbonio. Molte di esse si erano ridotte a pesare poco più di una palla da tennis, ma conservavano la forma e la dimensione originarie. Dato che queste «palle di ruggine» si sarebbero disgregate se esposte all'aria, le facemmo asciugare a 1060 gradi Celsius in atmosfera di idrogeno, ripristinando così il ferro ed eliminando i composti potenzialmente dannosi.

Il problema successivo consistette nel rimontare la nave. Fortunatamente lo scafo era stato costruito con un duplice sistema di perni in ferro e caviglie in legno, sicché la distruzione degli elementi metallici non aveva fatto sì che la struttura si sfasciasse. Il grosso dello scafo era quindi stato recuperato in un solo pezzo. Il cassero di poppa, tuttavia, si era spezzato e la serpe - la struttura sporgente a prua, sotto il bompresso - e altre parti si erano staccate. Di conseguenza fu necessario identificare e ricollocare molte migliaia di componenti strutturali che andavano da travi enormi a minuscoli pezzetti di legno: un rompicapo molto complesso da risolvere senza l'aiuto di un modello.

I restauratori dubitarono a lungo dell'opportunità di integrare le lacune della documentazione archeologica. In un primo tempo, data la quantità enorme di parti originali rimaste, essi riuscirono a eludere il problema, ma presto arrivarono a un punto in cui era necessario costruire strutture di connessione per poter utilizzare alcuni dei reperti conservati. Ormai, però, si era deciso che la cosa migliore era ricostruire la nave il più completamente possibile, anche perché erano pochissime le parti da realizzare da zero. La simmetria ci aiutò enormemente: dato che una nave ha due fianchi uguali, nella maggior parte dei casi riuscimmo a trovare una controparte in base alla quale ricostruire gli elementi mancanti. Naturalmente, le parti rifatte sono state evidenziate come tali.

Il restauro della *Vasa* ha rivelato molti aspetti inattesi delle costruzioni navali dell'epoca. Il sistema usato per dirigere la nave, per esempio, è stranamente laborioso. La barra del timone si estendeva per 11 metri dalla sommità della pala fino alla controbarra verticale, un'asta che sporgeva attraverso il ponte, collegata perpendicolarmente alla prima tramite un cuscinetto a rulli. Per fare accostare la nave, il timoniere doveva prima inclinare la controbarra e poi premerla verso il basso con tutto il suo peso attraverso il cuscinetto; questa azione combinata faceva muovere lateralmente la barra del timone. Per centrare quest'ultimo, doveva riportare la controbarra alla posizione iniziale attraverso il foro nel cuscinetto.

Sembra che il maestro d'ascia abbia seguito una tradizione che aveva un senso solo quando le navi erano più piccole. Nella *Vasa* la lunghezza eccessiva della barra limitava il movimento della pala del timone a 14 gradi, sette a sinistra e sette a destra. Questo fatto, insieme con la superficie relativamente piccola della pala, faceva sì che per governare la nave fosse necessario affidarsi soprattutto alla manovra delle vele.

La caratteristica più evidente della *Vasa* era la ricchissima decorazione, che tanto contrasta con il moderno concetto di subordinazione della forma alla funzione. Questa modificazione del gusto può essere esemplificata in modo addirittura più efficace dalla *Vasa* che non dalle testimonianze architettoniche, perché è più logico attendersi di trovare ornamenti in un edificio anziché in una nave da guerra; tuttavia il vascello è ampiamente ornato con sculture abbastanza vistose da attirare l'attenzione durante un passaggio ravvicinato. Gli scopi di questo apparato decorativo erano molteplici: incoraggiare gli alleati, intimidire i nemici, sostenere rivendicazioni e impressionare gli osservatori con simboli di potenza e gloria.

La decorazione della poppa, per esempio, è una celebrazione dei Vasa, la famiglia reale svedese da cui la nave prende il nome. In alto si vede un ritratto a bassorilievo del giovane Gustavo Adolfo, con la corona reale sorretta sopra il suo capo da due grifoni. Questa immagine è un'affermazione politica, in quanto contraddice la pretesa di Sigismondo III di Polonia - primo cugino di Gustavo Adolfo - al trono di Svezia, ed è rafforzata dal fregio sottostante, che reca le lettere «G A R S», acronimo di Gustavus Adolphus Rex Sueciae. Più in basso appare lo stemma nazionale svedese, e sotto di esso il cimiero della dinastia reale; Vasa è un antico termine svedese che significa «covone», e la relativa immagine appare nell'insegna.

Il colore era quasi totalmente scomparso dalle sculture, il cui legno imbevuto d'acqua aveva assunto una tinta marrone scuro quasi uniforme. A un esame più attento si scorgevano qua e là frammenti di foglia d'oro, a dimostrazione del fatto che molte delle sculture erano dorate. Eravamo in dubbio sugli altri colori, ma dal 1961 a oggi sono state messe a punto tecniche analitiche che ci hanno consentito di distinguerli; in particolare si è rivelata efficace la combinazione di microscopia elettronica e analisi con i raggi X. È ora chiaro che la *Vasa* era splendente di colori e oro come la pala d'altare di una chiesa barocca. Finora abbiamo utilizzato questa informazione solo per restaurare alcune sculture, in particolare il cimiero della famiglia Vasa.

La vita a bordo di una nave da guerra era dura, anche se il terzo decennio del Seicento vide poche battaglie navali. Ma anche il tempo di pace aveva i suoi orrori: oltre 400 uomini dovevano ammassarsi in uno spazio limitato. Le condizioni igieniche sulle navi dell'epoca erano terrificanti, e il cibo era cattivo già in porto e quasi immangiabile dopo essere rimasto per qualche settimana nella stiva. Marinai e ufficiali erano quindi esposti a un gran numero di infezioni.

Una nave naufragata è una sorta di microcosmo congelato nel tempo; è una testimonianza molto preferibile ai «ripostigli» di oggetti nascosti o ai corredi tombali, in quanto fornisce un'istantanea completa e involontaria di vita quotidiana. Le navi, le loro attrezzature e i loro carichi contengono una mole di informazioni sulla tecnologia e il commercio, le guerre e la diplomazia, la gente comune e lo spirito della loro epoca.

La *Vasa* non aveva a bordo l'equipaggio e l'attrezzatura al completo quando affondò; per esempio non c'erano i 300 fanti di marina che dovevano costituire il contingente militare imbarcato. Tuttavia sono state recuperate migliaia di oggetti personali; quelli in possesso dei marinai erano per lo più di legno o di terracotta, ma negli alloggi degli ufficiali, a poppa, sono stati trovati preziosi oggetti di vetro, porcellana, peltro e ottone, sia di fabbricazione svedese sia di importazione da altri paesi d'Europa. Una bottiglia di peltro conteneva ancora una bevanda alcolica a 66 gradi che sembra fosse arrak delle Indie Orientali; per esperienza personale, posso affermare che il liquore era ottimo.

Nel giugno 1990 il re di Svezia, Carlo XVI Gustavo, inaugurò il Museo Vasa a Stoccolma; qui si può visitare la nave ammiraglia circondata da memorie del mondo da cui ci è giunta: una testimonianza davvero unica della sua epoca.

BIBLIOGRAFIA

KVARNING LARS-ÅKE e OHRELIUS BENGT, *The Swedish Warship Wasa*, Macmillan, 1974.

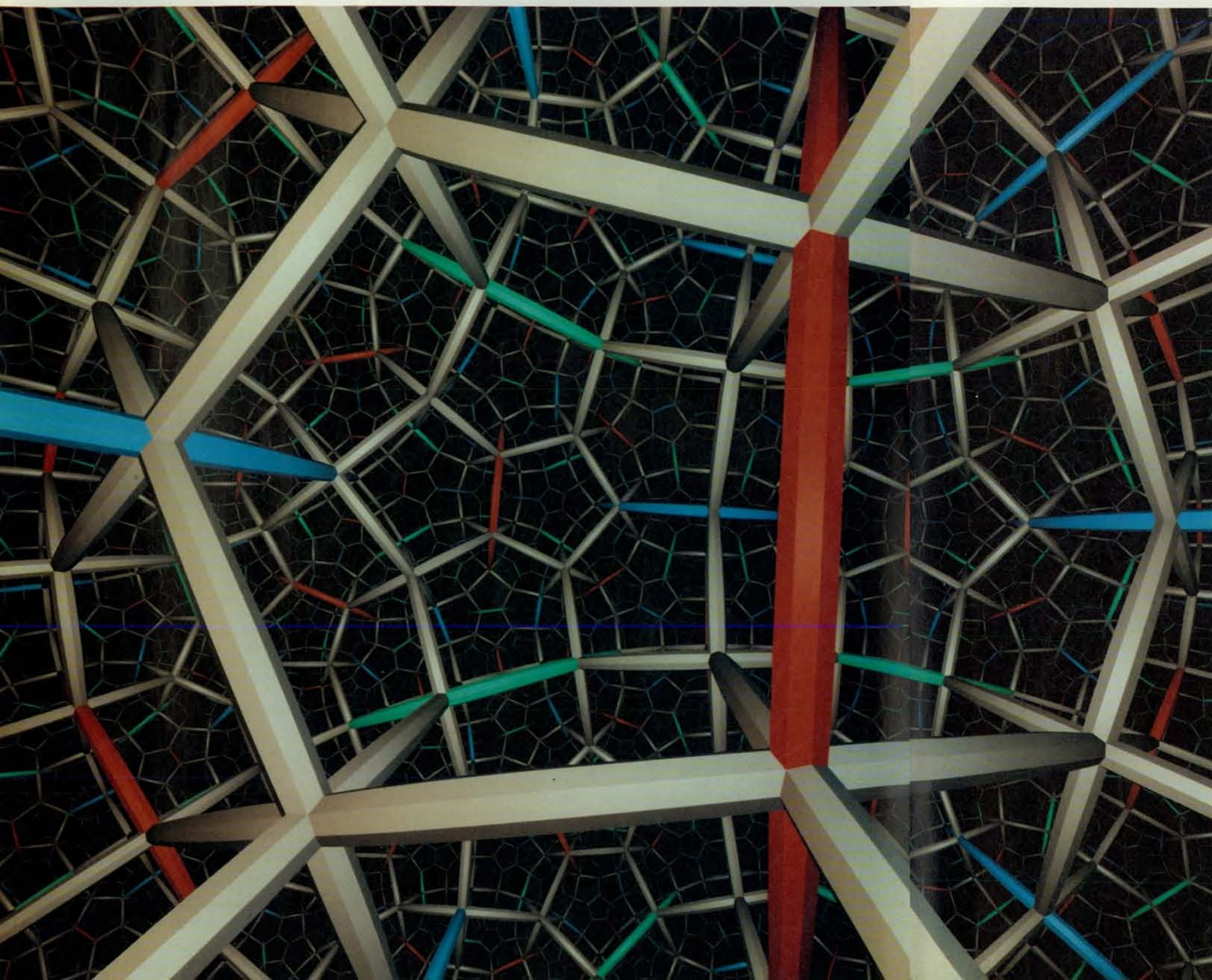
BORGENSTAM CURT e SANDSTRÖM ANDERS, *Why Wasa Capsized*, Wasa Studies n. 13, National Maritime Museum, 1984.

SOOP HANS, *The Power and the Glory: The Sculptures of the Warship Wasa*, Almqvist and Wiksell/Coronet Books, 1986.

LANDSTRÖM BJÖRN, *The Royal Warship Vasa*, Stenström Interpublishing, 1988.

ADRUP K. A., *The Vasa Museum: An Old Ship in a New House*, National Board of Public Building, 1990.

MATZ ERLING, *Vasa*, Vasa Museum, 1990.

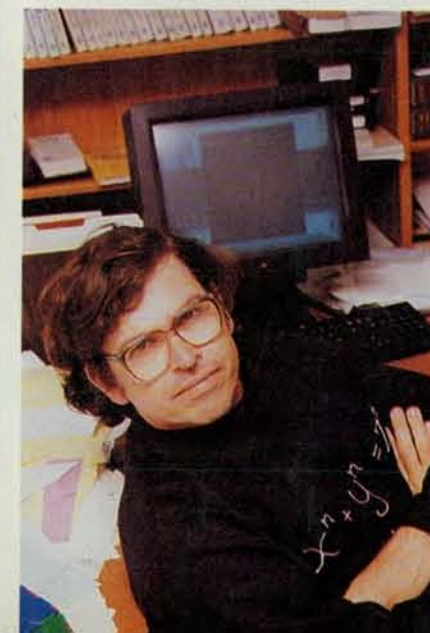
MORTE DELLA DIMOSTRAZIONE

Il calcolatore sta trasformando i metodi con cui i matematici scoprono, dimostrano e comunicano le loro idee, ma c'è da domandarsi se in questo nuovo mondo resti ancora spazio per la certezza assoluta

di John Horgan

La leggenda narra che quando Pitagora e i suoi discepoli scoprirono il celebre teorema, nel VI secolo a.C., sacrificarono un bove per festeggiare l'evento. E ne avevano ben donde, perché la relazione da essi scoperta tra i lati di un triangolo rettangolo valeva non in determinate circostanze, bensì sempre, che il triangolo fosse un taglio di seta, un appezzamento di terreno o un disegno su papiro: appariva una magia, un dono degli dèi. Non fa meraviglia che tanti pensatori, da Platone a Kant, abbiano reputato la matematica capace di offrirci le verità più pure che all'uomo sia concesso di conoscere.

Quest'opinione sembrò confermata nel giugno scorso, quando Andrew J. Wiles della Princeton University comunicò durante un congresso tenutosi all'Università di Cambridge di aver dimostrato l'ultimo teorema di Fermat. Si tratta di uno dei problemi più famosi della matematica: formulato più di 350 anni fa, le sue radici risalgono allo stesso Pitagora. Dato che in sala non si trovò alcuna vittima sacrificale, il pubblico manifestò il suo apprezzamento con un applauso.

Tuttavia ci si può chiedere se la dimostrazione dell'ultimo teorema di Fermat non sia stato l'estremo sussulto di una cultura morente. La matematica, che tra le discipline intellettuali è la più legata alla tradizione, sta subendo profonde trasformazioni. Per millenni i matematici hanno commisurato i loro progressi a ciò che si può dedurre tramite la dimostrazione, cioè una successione di passaggi logici che da una serie di assiomi porta a una conclusione irrefutabile. Ebbene, i dubbi che travagliano il pensiero odierno hanno ormai contaminato anche la matematica. Può darsi che i matematici siano prima o poi costretti ad accettare ciò che già molti scienziati e filosofi hanno ammesso, cioè che le loro asserzioni sono, nella migliore delle ipotesi, vere solo provvisoriamente, finché non se ne dimostri la falsità.

Quest'incertezza deriva in parte dalla crescente complessità della matematica. Spesso le dimostrazioni sono così lunghe e complicate che è difficile darne una valutazione: la dimostrazione di Wiles riempie ben 200 pagine, e gli esperti ritengono che sarebbe stata cinque volte più lunga se tutti i passaggi fossero stati esplicitati. Un osservatore ha affermato che solo un matematico su mille è in

La «dimostrazione video» qui a fianco illustra un teorema, dimostrato da W. P. Thurston del Mathematical Sciences Research Institute, che istituisce un profondo collegamento tra topologia e geometria. Il teorema asserisce che nello spazio circostante un nodo complesso (rappresentato in questa scena da un reticolo) vale una geometria «iperbolica», in cui le rette parallele divergono e i lati dei pentagoni formano angoli retti. Il video, generato al calcolatore e intitolato *Not Knot*, è stato prodotto dal Geometry Center dell'Università del Minnesota.

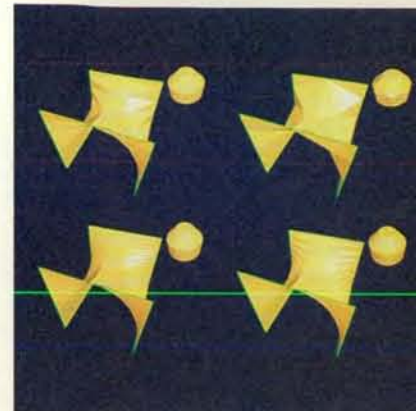
grado di valutare questa dimostrazione. Wiles è stato ritenuto credibile soprattutto grazie alla reputazione sua e di coloro sui cui risultati ha basato la dimostrazione. Alcuni matematici che non l'avevano ancora esaminata nei dettagli hanno tuttavia affermato che «sembra bellissima» e che «ha l'aria di essere vera».

Un altro catalizzatore di cambiamento è il calcolatore, che sta costringendo i matematici a riconsiderare la natura stessa della dimostrazione, e quindi della verità. Per ottenere certe dimostrazioni, negli ultimi anni si sono dovute eseguire masse enormi di calcoli, sicché nessun essere umano può verificare queste cosiddette dimostrazioni al calcolatore: solo altri calcolatori sono in grado di farlo. Di recente alcuni ricercatori hanno proposto una dimostrazione al calcolatore che fornisce solo una probabilità, e non la certezza, della verità, il che per certi matematici è una vera e propria incongruenza. Altri ancora stanno preparando «dimostrazioni video», nella speranza che siano più convincenti di pagine e pagine di formule.

Allo stesso tempo, alcuni matematici mettono in dubbio l'idea che le dimostrazioni formali siano il metro assoluto della verità. Benché nessuno sostenga di poter fare del tutto a meno delle dimostrazioni, alcuni esperti ritengono che sia più facile stabilire la validità di certe proposizioni confrontandole con esperimenti eseguiti al calcolatore o con fenomeni del mondo reale. «Penso che nei prossimi cinquant'anni l'importanza della dimostrazione in matematica diminuirà» dice Keith Devlin del Colby College, che cura una rubrica sui calcolatori per «Notices of the American Mathematical Society». «Saranno molti di più coloro che faranno matematica senza necessariamente fare dimostrazioni.»

Queste eresie vengono propugnate da potenti forze istituzionali. Da molti anni la National Science Foundation esorta i matematici a dedicarsi di più all'informatica e ad altri campi suscettibili di applicazioni. Anche alcuni luminari come Phillip A. Griffiths, direttore dell'Institute for Advanced Study di Princeton, e Michael Atiyah, insignito nel 1966 della Fields Medal (considerata il premio Nobel per la matematica) e direttore dell'Isaac Newton Institute for Mathematical Sciences di Cambridge, hanno invitato i matematici a uscire dalla loro torre d'avorio e a prendere contatto con il mondo reale. In un momento in cui finanziamenti e posti di lavoro sono scarsi, i giovani matematici non possono permettersi di ignorare queste esortazioni.

Naturalmente vi sono sacche di resistenza. Alcuni ricercatori deplorano con amarezza l'informatizzazione della loro disciplina e l'importanza crescente delle (oh, che parola volgare!) «applicazioni». Uno dei più attivi paladini della tradizione è Steven G. Krantz della Washington University, che in conferenze e articoli ha ammonito gli studenti a scegliere la matematica e non l'informatica, che secondo lui potrebbe essere una moda passeggera. L'anno scorso, ricorda Krantz, un funzionario della National Science Foundation visitò l'università e dichiarò che il suo ente non poteva più permettersi di finanziare ricerche matematiche che non fossero «finalizzate». «Potevamo ribellarci e dire che è un grave errore» brontola Krantz, «ma i matematici sono degli smidollati e ribellarsi non rientra nelle loro tradizioni.»

David Mumford della Harvard University, che è stato insignito della Fields Medal nel 1974 per ricerche di matematica pura e che ora studia la visione artificiale, ha scritto di recente che «nonostante le iperboli giornalistiche, le pressioni degli enti finanziatori e così via, per lo più i matematici puri considerano ancora i calcolatori come invasori e saccheggiatori del sacro suolo». L'anno scorso Mumford propose un corso in cui i docenti avrebbero dovuto insegnare agli studenti come si programma un calcolatore per risolvere problemi di analisi superiore. «Mi opposero il veto» ricorda «e non per le possibili proteste degli studenti, come mi aspettavo, ma perché metà dei miei colleghi non sapeva programmare!».

Jean E. Taylor, che studia matematica sperimentale alla Rutgers University, indaga sulle regole cui obbediscono le superfici minime esaminando oggetti sia reali, come le bolle di sapone, sia generati al calcolatore, come i cristalli idealizzati (*a sinistra*).

Questa situazione sta cambiando rapidamente, a giudicare dalle indicazioni fornite dal Geometry Center dell'Università del Minnesota. Fondato due anni fa, esso occupa il quinto piano di uno sfavillante poliedro di vetro e acciaio a Minneapolis e riceve due milioni di dollari l'anno dalla National Science Foundation, dal Department of Energy e dall'università. Tra i docenti fissi, che per lo più hanno anche incarichi esterni, vi sono alcuni dei più eminenti matematici del mondo.

Prendiamo una tipica giornata al Geometry Center: alcuni giovani ricercatori stanno ultimando un video che mostra come si può sminuzzare, torcere, stirare e infine rivoltare una sfera. In un'aula tre esperti in informatica di grandi università spiegano a una ventina di insegnanti di scuola superiore come allestire programmi grafici per l'insegnamento della matematica. Davanti a terminali NeXT color antracite, altri ricercatori meditano su sgargianti immagini di «ipercubi» quadridimensionali, di frattali vorticanti e di reticoli che si proiettano verso l'infinito. In giro non si vedono né carta né penne.

Di fronte a un terminale c'è David Ben-Zvi, un giovane ricercatore di Princeton che trascorre qui un periodo di sei mesi per studiare dinamica non lineare. Non è d'accordo con quei matematici che temono i calcolatori perché potrebbero allontanarli dai metodi che per tanto tempo si sono dimostrati utilissimi. «Hanno solo paura del cambiamento» dice garbatamente.

Il Geometry Center è una fucina della matematica sperimentale, in cui le ipotesi sono verificate tramite la rappresentazione grafica e l'uso del calcolatore. L'anno scorso alcuni membri del centro contribuirono alla fondazione di una rivista, «Experimental Mathematics», che illustra questi studi. «I metodi sperimentali non sono una novità in matematica» dice il direttore della rivista, David B. A. Epstein dell'Università di Warwick in Inghilterra, ricordando che Carl Friedrich Gauss e altri grandi matematici spesso effettuavano calcoli sperimentali prima di elaborare una dimostrazione formale. «La novità è che adesso sono diventati rispettabili.» Epstein ammette che non tutti i suoi colleghi sono d'accordo. «Uno di loro ha detto che la nostra rivista si dovrebbe chiamare "Rivista dei teoremi non dimostrati".»

Bolle di sapone e fusilli

Una studiosa che bene esemplifica il nuovo stile della matematica è Jean E. Taylor della Rutgers University. «Il principio secondo cui non si deve usare il calcolatore sarà sempre più estraneo alla nuova generazione» dice la Taylor, che da vent'anni studia le superfici minime (l'area o il volume più piccoli limitati da una curva o superficie). Forse le superfici minime più semplici ed eleganti che si trovino in natura sono le bolle e le pellicole di sapone. La Taylor ha sempre avuto un'inclinazione sperimentale: agli inizi della carriera, collaudava i suoi modelli teorici delle superfici minime immergendo anelli di fil di ferro in un secchio di acqua e sapone.

Oggi preferisce costruire modelli di bolle mediante un raffinato programma di grafica al calcolatore; inoltre è passata dalle bolle di sapone ai cristalli, che obbediscono a regole un po' più complicate sulle superfici minime. Insieme con Frederick J. Almgren di Princeton, Robert F. Almgren dell'Università di Chicago (che sono rispettivamente suo marito e suo figlio) e Andrew R. Roosen del National Institute of Standards and Technology, la Taylor sta cercando di simulare al calcolatore l'accrescimento di fiocchi di neve e altri cristalli. Collabora sempre più strettamente con fisici e specialisti di scienza dei materiali, offrendo loro teorie matematiche e tecniche di programmazione in cambio di informazioni sull'accrescimento dei cristalli reali.

Un altro esploratore del ciberspazio in cerca di nuove superfici minime è David A. Hoffman dell'Università del Massachusetts ad Amherst. Tra i suoi obiettivi preferiti vi sono i catenoidi e gli elicoidi, che assomigliano a fusilli e furono scoperti già nel Settecento. «Osservando le immagini di queste superfici al calcolatore» dice Hoffman, «ricaviamo un'enorme quantità di nuove idee».

Nel 1992 Hoffman, Fusheng Wei di Amherst ed Hermann Karcher dell'Università di Bonn ipotizzarono l'esistenza di una nuova classe di elicoidi, dotati di manici; solo dopo aver rappresentato questi elicoidi (i primi scoperti dal Settecento in poi) al calcolatore, essi riuscirono a fornire una dimostrazione formale della loro esistenza. «Se non avessimo potuto vedere una figura grosso modo corrispondente a ciò che immaginavamo, non saremmo mai riusciti nell'impresa» dice Hoffman.

Il settore della matematica sperimentale che nell'ultimo decennio ha fatto la parte del leone è la cosiddetta dinamica

Uno splendido anacronismo?

Quanti considerano la matematica sperimentale e le dimostrazioni al calcolatore un obbrobrio più che un'innovazione hanno un motivo particolare per rallegrarsi della dimostrazione dell'ultimo teorema di Fermat da parte di Andrew J. Wiles della Princeton University. L'impresa di Wiles è un trionfo della tradizione e si pone controcorrente rispetto a tutta la matematica moderna.

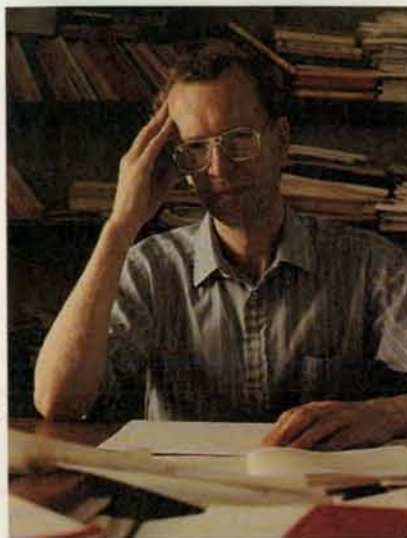
Wiles, strenuo sostenitore della matematica per la matematica, afferma: «Non vorrei certo vedere la matematica ridotta a schiava delle applicazioni, perché questo non sarebbe neppure nell'interesse delle applicazioni».

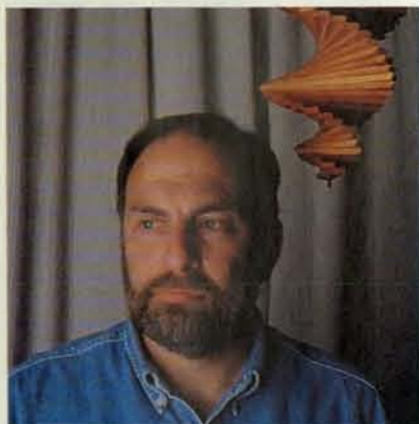
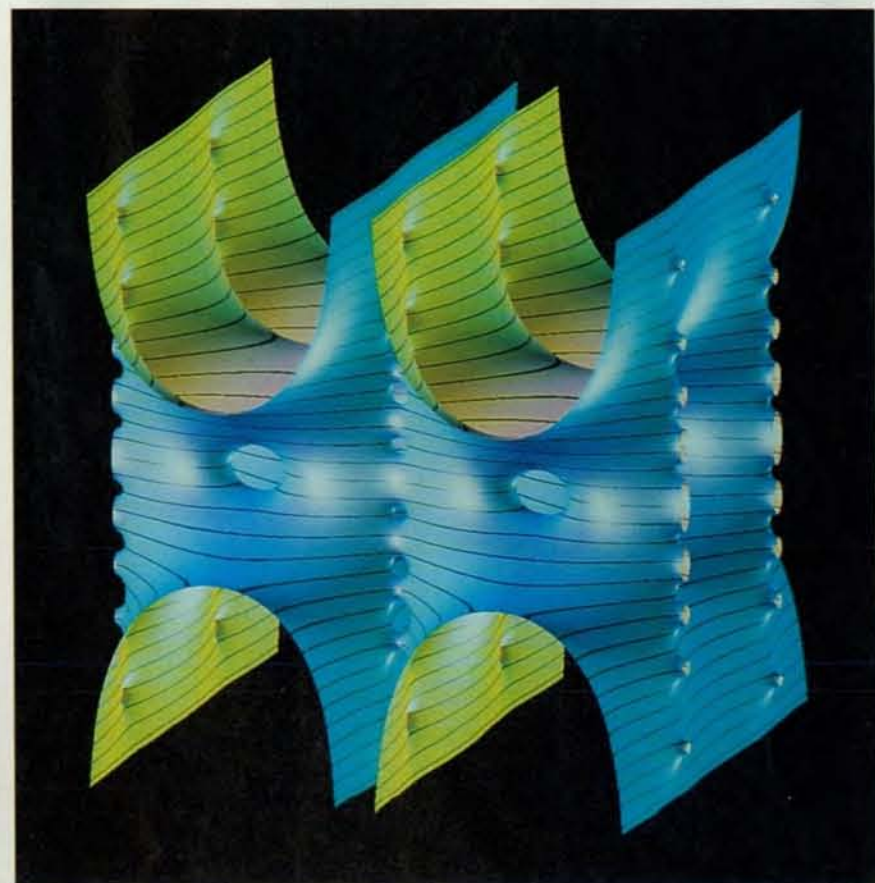
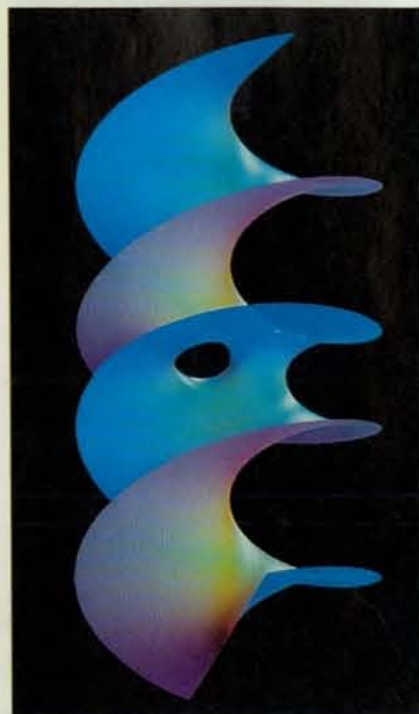
Il problema da lui risolto, posto più di 350 anni or sono dal matematico francese Pierre de Fermat, è uno splendido rompicapo di matematica pura. Fermat dichiarò di aver trovato una dimostrazione della proposizione seguente: per tutti i valori di n maggiori di 2, l'equazione $x^n + y^n = z^n$ non ammette soluzioni intere. I tentativi di trovarne la dimostrazione, mai rivelata da Fermat, hanno contribuito alla costruzione della teoria dei numeri moderna (lo studio dei numeri interi), che di recente è risultata utile in crittografia. Eppure «è molto improbabile che di per sé l'ultimo teorema di Fermat abbia qualche applicazione» dice Wiles.

Benché da tempo gli enti finanziatori invitino i matematici a collaborare sia tra loro sia con altri scienziati, Wiles ha lavorato pressoché da solo per sette anni. Soltanto verso la fine della ricerca ha comunicato le proprie idee ad alcuni colleghi.

La dimostrazione di Wiles ha in sostanza la stessa forma deduttiva che hanno i teoremi classici della geometria euclidea. Non contiene calcoli e rivendica la verità assoluta (non statistica). Wiles non ha usato il calcolatore nemmeno per dare una rappresentazione grafica alle sue idee, per effettuare calcoli o per scrivere la sua memoria, che una segretaria ha battuto a macchina dagli appunti manoscritti.

Egli ammette che verificare una congettura con il calcolatore può essere utile. Negli anni settanta certe verifiche al calcolatore indicarono che una cervellotica ipotesi, la congettura di Taniyama, poteva essere vera. Le ricerche nate da queste verifiche sono alla base della dimostrazione di Wiles. Tuttavia egli non crede che si darà la pena di imparare a fare ricerca col calcolatore. «È un'abilità a parte» spiega «e se s'investe tanto tempo in un'abilità a parte, è probabile che ci si lasci sviare dal vero lavoro sul problema».

Wiles respinge la possibilità che possa esistere un numero finito di verità accessibili alle forme tradizionali d'indagine. «Mi oppongo con forza all'idea che i buoni teoremi comincino a scarseggiare» dice. «Credo anzi che abbiamo appena scalfito la superficie.»

non lineare, più nota al pubblico con l'improprio pseudonimo di caos. In genere i sistemi non lineari sono retti da insiemi di regole semplici che, grazie alla retroazione e ad altri effetti analoghi, danno luogo a fenomeni complessi. I sistemi non lineari, beninteso, furono studiati anche prima dell'avvento dell'informatica, ma ora i calcolatori consentono di esplorare questi sistemi e di osservarne l'evoluzione con metodi di cui Henri Poincaré e gli altri pionieri di questa branca della matematica non potevano disporre.

Gli automi cellulari, che dividono lo schermo del calcolatore in un insieme di celle analoghe agli elementi d'immagine, o pixel, illustrano in maniera lampante i principi della non linearità. In generale il colore, o «stato», di ciascuna cella è determinato da quello delle celle vicine; se lo stato di una singola cella cambia, una cascata di cambiamenti si diffonde in tutto il sistema.

Uno degli automi cellulari più famosi è quello inventato da John H. Conway di Princeton all'inizio degli anni settanta. Conway ha dimostrato che il suo automa, battezzato «Vita», è indecidibile: cioè non si può stabilire se le sue configurazioni varino all'infinito oppure se prima o poi siano destinate a ripetersi. Alcuni scienziati hanno sfruttato gli automi cellulari per studiare l'origine e l'evoluzione della vita; l'informatico e fisico Edward Fredkin dell'Università di Boston arriva addirittura a sostenere che l'universo nel suo complesso sia un automa cellulare.

Ancor più famoso è l'insieme di Mandelbrot, la cui immagine è diventata un simbolo di tutti gli studi sul caos da quando Benoit B. Mandelbrot, del Thomas J. Watson Research Center della IBM, lo ha reso popolare nei primi anni ottanta. Questo insieme scaturisce da una semplice equazione comprendente un termine complesso (cioè un termine che contiene la radice quadrata di un numero negativo); l'equazione produce soluzioni che vengono poi iterate, ossia reintrodotte nell'equazione.

La teoria di questo insieme era stata formulata più di settant'anni fa da due matematici francesi, Gaston Julia e Pierre Fatou, ma solo il calcolatore ne ha rivelato a tutti la barocca bellezza. Quando viene rappresentato su uno schermo, l'insieme di Mandelbrot genera un'immagine che è stata paragonata a un cuore carcinomatoso, a un pollo bruciacchiato o a un pupazzo di neve bitorzoluto. L'immagine è un frattale: i suoi margini sfumati sono di lunghezza infinita e vi compaiono configurazioni ricorrenti a scale diverse.

Oggi si stanno studiando insiemi che sono simili all'insieme di Mandelbrot, ma si sviluppano in quattro dimensioni. «Le complicazioni che s'incontrano in questo caso sono simili a quelle che ci si trova a dover affrontare anche in molte altre scienze» dice John Milnor, della State University of New York a Stony Brook. Milnor sta attualmente tentando di analizzare le proprietà dell'insieme a quattro dimensioni esaminandone sezioni bidimensionali generate al calcolatore. I risultati preliminari da lui ottenuti sono stati pubblicati lo scorso anno nell'apertura del primo numero di «Experimental Mathematics». Milnor, che è stato insignito della Fields Medal nel 1962, racconta come ogni tanto si cimentasse in esperimenti al calcolatore anche ai tempi delle schede perforate, ma «era un procedimento penoso. Ora è molto più facile».

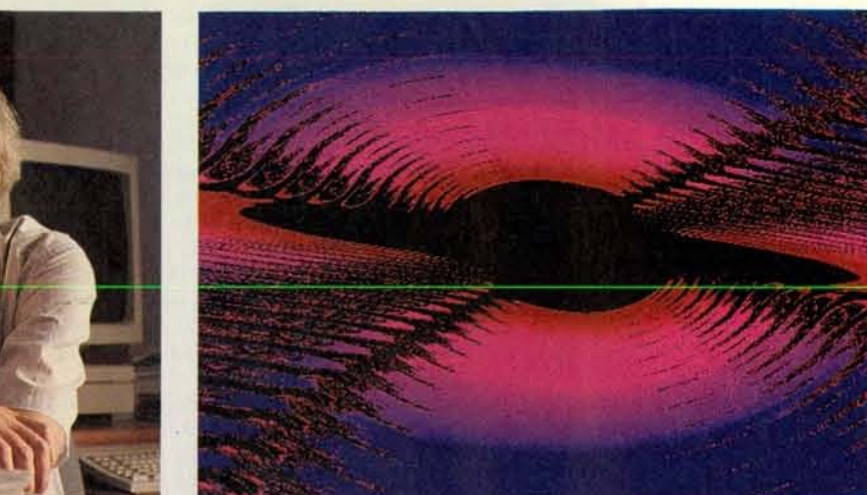
Questo elicoide forato (*in basso a sinistra*) è stato scoperto l'anno scorso da David A. Hoffman e colleghi dell'Università del Massachusetts ad Amherst servendosi della grafica al calcolatore. Di recente Edward C. Thayer, del gruppo di Hoffman, ha individuato una struttura (*qui sotto*) che simula la configurazione di certi polimeri.

Paesaggi alieni appaiono generati da un calcolatore che produce sezioni di un oggetto quadridimensionale simile al ben noto insieme di Mandelbrot. John Milnor della State University of New York a Stony Brook studia analoghe immagini bidimensionali per comprendere le proprietà di questo complesso oggetto matematico.

La popolarità della matematica basata sulla grafica ha provocato una reazione. Quattro anni fa, in «Mathematical Intelligencer», Krantz ha denunciato il fatto che «in certi ambienti è più facile ottenere fondi per comprare macchine che producano immagini di frattali che non per studiare geometria algebrica».

Un ammonimento di carattere più generale sulla matematica «speculativa» è comparso nel luglio scorso in «Bulletin of the American Mathematical Society». Gli autori, Arthur Jaffe di Harvard e Frank S. Quinn del Virginia Polytechnic Institute, sostengono che nello stabilire la verità gli esperimenti al calcolatore e la corrispondenza con fenomeni naturali non possano sostituirsi alle dimostrazioni. Scrivono Jaffe e Quinn: «Di quando in quando, in seno alla comunità dei matematici, vi sono stati gruppi e individui che hanno tentato di attenuare il rigore delle dimostrazioni; i risultati sono stati contrastanti e talvolta disastrosi».

Anche i matematici che sfruttano la grafica al calcolatore e altre tecniche sperimentali per lo più concordano sul fatto che vedere non basta per credere e che le dimostrazioni sono ancora necessarie per verificare le congetture ottenute con il calcolo. «Io credo che per troppo tempo i matematici siano stati legati al passato, ma questo non implica che le dimostrazioni non siano importanti» dice la Taylor. Hoffman difende la dimostrazione tradizionale con forza ancora maggiore: «Le dimostrazioni sono gli unici strumenti di laboratorio che i matematici posseggano, e c'è il pericolo che stiano per essere gettati via». Benché la grafica al calcolatore sia «straordinaria e meravigliosa» aggiunge, «negli anni sessanta anche gli stupefacenti erano straordinari e meravigliosi, ma alcune persone ci rimisero la pelle».

In effetti gli informatici esperti sanno meglio di chiunque altro che gli esperimenti al calcolatore (che siano basati sulla grafica o sul calcolo numerico) possono essere ingannevoli. Un esempio ammonitore riguarda l'ipotesi di Riemann, una famosa congettura su come si distribuiscano i numeri primi nel tendere all'infinito. Avanzata da Bernhard Riemann più di cent'anni fa, questa ipotesi è considerata uno dei più importanti problemi non risolti della matematica.

Franz Mertens, un matematico dell'epoca di Riemann, avanzò una congettura analoga relativa ai numeri interi positivi; se si fosse rivelata vera, essa avrebbe costituito un valido argomento a favore dell'ipotesi di Riemann. All'inizio degli anni ottanta si era stabilito, con l'uso del calcolatore, che in effetti l'ipotesi di Mertens valeva almeno per i primi dieci miliardi di interi; nel 1984, ulteriori calcoli dimostrarono tuttavia che a un certo punto - per numeri dell'ordine di $10^{(10^{70})}$ - la distribuzione prevista da Mertens scompare.

Un potenziale svantaggio dei calcolatori è che tutte le loro operazioni si basano sulla manipolazione di numeri interi discreti (per la precisione, dei numeri zero e uno); i numeri reali, come π o la radice di 2, possono essere solo approssimati. Chiunque abbia un po' di esperienza con le funzioni di arrotondamento di una semplice calcolatrice tascabile può indurla facilmente a fornire risultati errati. Programmi più raffinati possono introdurre errori più complicati ed elusivi. Nel 1991 David R. Stoutemyer, specialista di software dell'Università di Hawaii, illustrò 18 esperimenti di algebra che, eseguiti con gli ordinari programmi per la matematica, fornivano risultati sbagliati.

Stephen Smale dell'Università della California a Berkeley, insignito della Fields Medal nel 1966, ha cercato di dare fondamenta più solide al calcolo ma-

tematico, o almeno di individuare la posizione e le dimensioni delle crepe che serpeggiano in queste fondamenta. Insieme con Lenore Blum del Mathematical Sciences Research Institute di Berkeley e Michael Shub della IBM, Smale ha allestito un modello teorico di calcolatore capace di elaborare numeri reali anziché semplici interi.

Smale e la Blum hanno di recente concluso che l'insieme di Mandelbrot è, in senso tecnico, non computabile: cioè non si può stabilire con certezza se un punto arbitrario del piano complesso si trovi all'interno o all'esterno della sua irta frontiera. Secondo Smale, questo risultato dimostra che, quando si estrapolano i risultati di esperimenti al calcolatore, è necessaria molta attenzione.

Stephen Wolfram, fisico matematico dell'Università dell'Illinois, non condivide queste preoccupazioni. Egli è l'inventore di *Mathematica*, che da quando è stato immesso sul mercato, cinque anni fa, è diventato il più diffuso software per la matematica. Egli riconosce che «in effetti nella matematica sperimentale ci sono dei trabocchetti. Come accade per tutti gli esperimenti, può capitare di sbagliare». Ma sottolinea che gli esperimenti al calcolatore, se compiuti e analizzati con intelligenza, possono fornire più risultati del vecchio metodo basato su congettura e dimostrazione. «In ogni altro campo della scienza gli sperimentatori sono molto più numerosi dei teorici» dice Wolfram «e credo che anche in matematica prevarrà sempre più questa tendenza».

Wolfram sostiene che «l'ossessione della dimostrazione» ha impedito ai matematici di scoprire i nuovi vasti domini dei fenomeni accessibili al calcolatore. E aggiunge che anche i più intrepidi matematici sperimentali per lo più non «osano abbastanza». «Della matematica prendono i problemi esistenti e li studiano. Non fanno altro che aggiungere qualche ghirigoro in cima a una struttura gigantesca.»

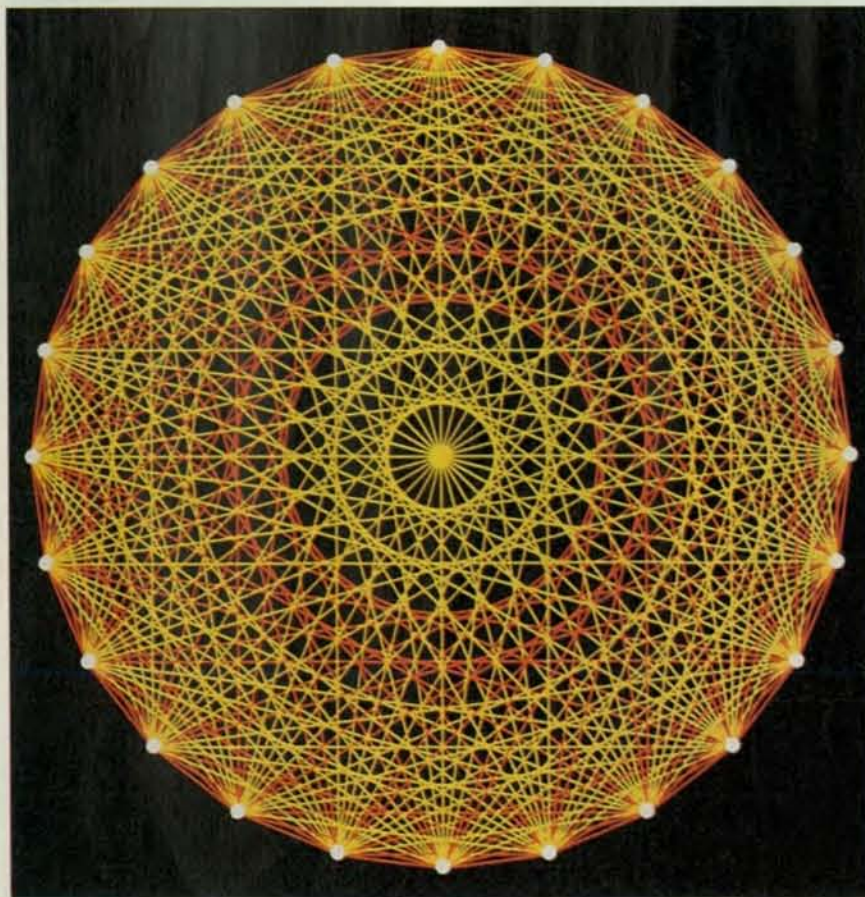
I matematici potrebbero forse essere d'accordo, ma con qualche riserva. Anche Conway è affascinato dagli automi cellulari, eppure sostiene che la carriera di Wolfram - come pure il suo disprezzo per la dimostrazione - indica che egli non è un matematico autentico. «Di solito i matematici puri non fondano società e non affrontano il mondo in modo così aggressivo. Noi preferiamo stare nella nostra torre d'avorio a riflettere.»

Più difficile è per i puristi ignorare la voce di un altro fervido sostenitore della sperimentazione e dell'uso dei calcolatori in matematica, William P. Thurston. Direttore del Mathematical Sciences Research Institute di Berkeley e condirettore (con Albert Marden dell'Università del Minnesota) del Geometry Center, Thurston vanta credenziali impeccabili: a metà degli anni settanta evidenziò la possibilità di una profonda connessione tra due branche separate della matematica, la topologia e la geometria, e per queste ricerche nel 1982 gli fu assegnata la Fields Medal.

Thurston crede fermamente che le verità matematiche siano scoperte e non inventate, ma per ciò che riguarda la dimostrazione non si rifà tanto a Platone quanto a Thomas S. Kuhn, il quale, in *La struttura delle rivoluzioni scientifiche*, sosteneva che le teorie scientifiche vengano accettate per ragioni sociali e non perché siano oggettivamente «vere». «Che in linea di principio la matematica si riduca alle dimostrazioni formali è un'idea senza solide basi» tipica di questo secolo, afferma Thurston. E aggiunge: «In pratica i matematici dimostrano i teoremi in un contesto sociale. È un corpo di conoscenze e di tecniche soggetto a condizionamenti sociali».

Thurston osserva che più di sessant'anni fa il logico Kurt Gödel dimostrò con il suo teorema di incompletezza che «è impossibile codificare la matematica». Qualunque insieme di assiomi fornisce enunciati che, pur essendo evidentemente veri, non possono essere dimostrati partendo da quegli assiomi. Ancor prima Bertrand Russell aveva fatto notare che la teoria degli insiemi, che sta alla base di gran parte della matematica, pullula di contraddizioni logiche legate al problema dell'autoreferenza (che è illustrato dall'enunciato autocontraddittorio «Questa proposizione è falsa»). «La teoria degli insiemi è basata su garbate bugie, cose su cui conveniamo anche se sappiamo che non sono vere» dice Thurston. «Per certi versi i fondamenti della matematica hanno qualcosa di irreale.»

Secondo Thurston le magagne si celano più facilmente nelle dimostrazioni altamente formali che in quelle dove ci si richiama a un livello di comprensione più intuitivo. In particolare è affascinato dalla capacità della grafica al calcolatore di comunicare concetti matematici astratti anche a persone estranee alla cerchia degli specialisti. Due anni fa, per sua sollecitazione, il Geometry Center allestì con l'aiuto del calcolatore una «dimostrazione video», *Not Knot*, che illustra una congettura rivoluzionaria da lui stesso dimostrata una decina d'anni fa. Thurston ricorda con orgoglio che il gruppo rock dei Grateful Dead ha proiettato il video durante i propri concerti.

Il problema del ricevimento è stato risolto da Stanislaw P. Radziszowski e Brendan D. McKay dopo lunghissimi calcoli. Essi hanno dimostrato che occorrono almeno 25 invitati per garantire che quattro di essi si conoscano tutti oppure che cinque siano tra loro estranei. Da questo diagramma, dove i segmenti rossi collegano gli amici e quelli gialli gli estranei, si vede che un gruppo di 24 invitati non soddisfa l'enunciato.

Matematici di silicio

La progressiva invasione della matematica da parte dei calcolatori ha riacceso un'antica disputa: la matematica può essere completamente automatizzata? I grandi matematici del prossimo secolo saranno di silicio?

In effetti gli informatici lavorano da decenni a programmi che generano congetture e dimostrazioni matematiche. Verso la fine degli anni cinquanta il guru dell'intelligenza artificiale Marvin Minsky mostrò come un calcolatore potesse «riscoprire» alcuni teoremi fondamentali della geometria euclidea. Negli anni settanta Douglas Lenat, ex studente di Minsky, presentò un programma che formulava teoremi di geometria ancora più avanzati. Gli scettici obiettarono che in realtà i risultati erano già contenuti nel programma originale.

Una decina d'anni fa l'informatico e imprenditore Edward Fredkin cercò di riaccendere il declinante interesse per la matematica eseguita dal calcolatore istituendo quello che poi fu chiamato premio Leibniz. Questo premio di 100 000 dollari, gestito dalla Carnegie Mellon University, spetta al primo programma per calcolatore capace di formulare un teorema che abbia un «profondo effetto» sulla matematica.

Alcuni esperti di quello che viene chiamato ragionamento automatico pensano di avere ormai diritto al premio. Uno di essi è Larry Wos dell'Argonne National Laboratory, direttore di «Journal of Automated Reasoning», il quale sostiene che il programma da lui allestito ha risolto problemi di matematica e di logica «che avevano sfidato i ricercatori per anni». Un altro è Siemeon Fajtlowicz dell'Università di Houston, inventore di un programma, «Graffiti», che ha proposto «migliaia» di congetture di teoria dei grafi.

Secondo David Mumford della Harvard University, che fa parte della giuria del premio, nessuno di questi risultati soddisfa neppure lontanamente il criterio del «profondo effetto». Se gli si chiede quando, secondo lui, verrà assegnato il premio, Mumford risponde: «Non ora, e neppure fra cent'anni».

Alcuni osservatori ritengono però che prima o poi i calcolatori supereranno le nostre capacità matematiche. In fin dei conti, osserva Ronald L. Graham degli AT&T Bell Laboratories, «noi non siamo molto adatti a elucubrare sul continuo spazio-temporale o sull'ipotesi di Riemann. Siamo progettati per raccogliere bacche o per evitare di essere mangiati».

Altri si schierano con il fisico matematico Roger Penrose dell'Università di Oxford, il quale nel suo libro *La mente nuova dell'imperatore* sostiene che i calcolatori non potranno mai sostituire i matematici. La sua argomentazione si rifà alla teoria quantistica e al teorema di incompletezza di Gödel, ma forse Penrose è più convincente quando discute la propria esperienza personale. Secondo lui, ai livelli più alti, la matematica è un'arte, un atto creativo che non può essere ridotto alla logica più di quanto sia possibile ricondurvi *Re Lear* o la *Quinta* di Beethoven.

Che poi gli ammiratori dei Grateful Dead afferrino completamente la sostanza del video (che riguarda il comportamento in uno spazio non euclideo «iperbolico» di oggetti matematici detti varietà tridimensionali) è un altro paio di maniche. Thurston ammette che il video è difficile da comprendere fino in fondo per i non matematici, e anche per alcuni specialisti, ma ciò non lo sgomenta. Il Geometry Center sta infatti allestendo un video su un altro dei suoi teoremi, che mostra come si possa rivoltare una sfera come un guanto. L'autunno scorso, inoltre, Thurston ha organizzato un convegno per discutere in che modo la realtà virtuale e altre tecnologie avanzate possano essere utili alla visualizzazione matematica.

È paradossale il fatto che i calcolatori abbiano catalizzato una controtendenza per cui la verità viene ottenuta a spese della comprensibilità. Nel 1976 Kenneth Appel e Wolfgang Haken dell'Università dell'Illinois annunciarono di aver dimostrato la congettura dei quattro colori, secondo la quale bastano quattro colori diversi per disegnare una carta geografica anche infinita in cui due paesi indicati con lo stesso colore non siano mai confinanti. Per certi versi la dimostrazione di Appel e Haken era tradizionale, cioè consisteva in una serie di passaggi logici e concatenati che portavano a una conclusione; e la conclusione era che la congettura poteva essere ricondotta alla verifica del comportamento di circa 2000 carte diverse.

Poiché eseguire a mano questo controllo sarebbe risultato proibitivo, Appel e Haken programmarono un calcolatore e gli affidarono la verifica. Dopo un migliaio di ore di calcoli, la macchina concluse che le 2000 carte geografiche si comportavano come previsto: dunque la congettura dei quattro colori era vera.

Il problema del ricevimento

In seguito sono state effettuate altre dimostrazioni con l'ausilio del calcolatore. Proprio quest'anno Stanislaw P. Radziszowski del Rochester Institute of Technology e Brendan D. McKay dell'Australian National University di Canberra hanno annunciato di aver dimostrato il cosiddetto problema del ricevimento. Questo problema, scaturito dalle ricerche sulla teoria degli insiemi compiute negli anni venti dal matematico britannico Frank P. Ramsey, può essere enunciato con una domanda sui rapporti intercorrenti tra gli ospiti di un ricevimento: qual è il minimo numero di persone che devono essere invitate per essere certi che almeno x ospiti si conoscano tutti o che almeno y ospiti siano tra loro estranei? Questo numero è detto numero di Ramsey.

Dimostrazioni ottenute in precedenza avevano stabilito che occorrono almeno 18 ospiti per garantire che vi siano quattro conoscenti o quattro estranei; con la loro dimostrazione Radziszowski e McKay hanno stabilito che il numero di Ramsey per quattro conoscenti o cinque estranei è 25. Gli amanti della mondanità ci pensino due volte prima di tentare il calcolo del numero di Ramsey per x e y più grandi: secondo le stime di Radziszowski e McKay, la dimostrazione ha richiesto l'equivalente di undici anni di calcoli con una normale macchina da tavolo. Radziszowski ritiene che, per i problemi di matematica pura, questo sia un primato.

Il significato di questa impresa è stato discusso in una sede inconsueta, la rubrica delle risposte ai lettori curata da Ann Landers su un quotidiano. In giugno un lettore scrisse alla Landers sostenendo che le risorse impiegate per risolvere il problema del ricevimento avrebbero dovuto essere usate per aiutare i «bambini affamati dei paesi devastati dalla guerra». Alcuni matematici sollevano un'altra obiezione contro le dimostrazioni assistite dal calcolatore. «Io non credo nelle dimostrazioni fatte dal calcolatore» dice Pierre Deligne dell'Institute for Advanced Study, specialista di geometria algebrica e insignito della Fields Medal nel 1978. «In un certo senso io sono molto egocentrico. Credo in una dimostrazione se la capisco, se è chiara.» Benché riconosca che le persone possano commettere errori, Deligne aggiunge: «Anche un calcolatore commette errori, e sono molto più difficili da scoprire».

Altri adottano un atteggiamento più pragmatico e sostengono che stabilire la verità è più importante che dare ai matematici una soddisfazione estetica, soprattutto se si vuole che un risultato trovi applicazioni. I sostenitori di questo punto di vista, che in genere sono informatici, osservano che anche le dimostrazioni tradizionali sono tutt'altro che esenti

Bolle di sapone: effimere e complesse

Tra i metodi della matematica sperimentale cui si accenna nell'articolo di John Horgan vi sono le bolle di sapone. Alle notevoli proprietà geometriche e fisiche di queste lucide sfere hanno da secoli dedicato la loro attenzione scienziati anche illustri, tra cui Isaac Newton, Robert Boyle e Robert Hooke. La proprietà matematica più importante delle bolle è senz'altro quella di minimo: dato un certo volume (quello dell'aria soffiata), la più piccola superficie capace di racchiuderlo è la sfera, forma che la bolla assume quasi per magia risolvendo questo classico problema matematico in un caso semplicissimo. Vi sono anche casi molto più complessi, che furono affrontati sistematicamente soprattutto da Joseph-Antoine-Ferdinand Plateau (1801-1883).

Plateau considerava una curva qualsiasi e, utilizzando proprio le lamine saponate, cercava, tra le superfici aventi come bordo la curva assegnata, quella di area minima. È la tensione superficiale (cioè la forza attrattiva tra le molecole) che «obbliga» una bolla o una lamina saponata ad assumere, all'equilibrio, una superficie che abbia la minima area compatibile coi vincoli imposti (il contorno). Plateau (che per ironia della sorte divenne cieco, ma continuò a studiare le bolle con l'aiuto di un fedele assistente) scoprì in questo modo che le forme delle lamine rette da telai metallici di svariatissime fogge (tetraedrica, dodecaedrica, elicoidale e così via) non sono casuali, ma obbediscono a regole precise.

Anche se oggi il calcolatore consente di studiare i problemi di minimo, e altri problemi connessi, con l'aiuto dell'eidomatica, le bolle e le lamine di sapone, oltre a esercitare il loro fascino di sempre, restano strumenti validi e praticissimi per suggerire, almeno in certi casi, utili soluzioni sperimentali. Per esempio, per il problema della rete di lunghezza minima - si pensi al solito commesso viaggiatore, che deve visitare tutti i suoi clienti, ma vuole percorrere la strada più breve (si veda «Le Scienze» n. 247, marzo 1989, p. 81) - si possono ottenere, in casi non troppo complessi, soluzioni immediate con un semplicissimo telaio, che a buon diritto si può chiamare «calcolatore a lamine di sapone».

Segnaliamo il volume che a queste eteree creature ha dedicato il matematico Michele Emmer (*Bolle di sapone*, La Nuova Italia, Firenze 1991, pp. 144, L. 120 000). Il libro, di grande interesse e splendidamente illustrato, sembra proprio ispirato al detto di Lord Kelvin: «Fate una bolla di sapone e osservatela: potreste passare tutta la vita a studiarla». Oltre a offrirci un'abbondante iconografia che documenta la presenza delle bolle nelle arti figurative dal Rinascimento in poi, l'autore ci guida alla scoperta di un mondo affascinante e sorprendente, in cui s'intrecciano letteratura e fisica, architettura e biologia, chimica, matematica e pittura. Un cenno a parte merita il capitolo dedicato alle affinità strutturali tra bolle di sapone e radiolari, minuscoli componenti del plancton. Questo capitolo è illustrato da alcuni dei disegni del naturalista Ernst Haeckel per uno dei cinquanta volumi di dati e osservazioni raccolti durante la spedizione della corvetta inglese *H.M.S. Challenger* tra il 1872 e il 1876. (*Giuseppe O. Longo*)

Telai a forma di tetraedro, di prisma regolare e di ottaedro tratti dal film di Michele Emmer *Soap bubbles*.

da errori. All'inizio del secolo quasi tutti i teoremi erano abbastanza brevi da poter essere letti da cima a fondo in una sola volta ed erano opera di un unico autore. Oggi le dimostrazioni occupano spesso centinaia di pagine e sono così complicate che possono passare anni prima che siano confermate da altri.

Tra le dimostrazioni tradizionali, il primato spetta attualmente alla classificazione dei gruppi semplici finiti, terminata nei primi anni ottanta. (Un gruppo è un insieme di elementi, per esempio gli interi, con un'operazione, per esempio l'addizione, che, dati due elementi, ne fornisce un terzo.) La dimostrazione consiste in circa 500 articoli, scritti da più di cento ricercatori, per un totale di quasi 15 000 pagine. Si è affermato che l'unica persona che aveva capito la dimostrazione nella sua interezza era il responsabile del progetto, Daniel Gorenstein della Rutgers University. Gorenstein è morto l'anno scorso.

Anche dimostrazioni molto più brevi possono sollevare dubbi. Tre anni fa Wu-Yi Hsiang di Berkeley annunciò di aver dimostrato una vecchia congettura secondo cui, per stipare in un volume assegnato il massimo numero di sfere, bisogna impilarle come palle di cannone. Oggi alcuni scettici sono convinti che nelle 100 pagine della dimostrazione si annidi un errore fatale; altri sono ugualmente certi che la dimostrazione sia sostanzialmente corretta.

In effetti, secondo alcuni specialisti di informatica, per migliorare l'affidabilità bisogna intensificare e non ridurre il ricorso al calcolatore. Robert S. Boyer dell'Università del Texas ad Austin ha diretto il Progetto QED, un tentativo di comprimere tutto il dilagante corpus della matematica moderna in un'unica base di dati, la cui coerenza potrà essere verificata mediante «controllori automatici di dimostrazioni».

Stando al manifesto del progetto, gli utenti di questa base di dati potranno «esaminare la totalità delle conoscenze matematiche per ricavarne i risultati a cui sono interessati e, ricorrendo agli strumenti del sistema QED, sfruttare questi risultati con sicurezza e fiducia, senza bisogno di comprenderne i dettagli e neppure i fondamenti ultimi». Il sistema QED, proclama con una certa magniloquenza il manifesto, potrà anche «costituire un antidoto agli effetti degenerativi del relativismo culturale e del nichilismo» e, presumibilmente, proteggere la matematica dalla tentazione del tutto umana di cedere alle mode.

Il dibattito sulle dimostrazioni al calcolatore si è intensificato di recente con l'avvento di una tecnica che offre non la certezza della verità, ma solo una probabilità statistica. Queste dimostrazioni sfruttano metodi simili a quelli su cui si basano i codici per la correzione degli errori, i quali, rendendo i messaggi trasmessi altamente ridondanti, li proteggono dal rumore e da altri effetti dannosi. La dimostrazione deve per prima cosa essere espressa in tutti i particolari secondo una forma rigorosa di logica matematica. Poi la logica subisce un'ulteriore trasformazione, l'aritmetizzazione, in cui «and», «or» e altre funzioni sono tradotte in operazioni aritmetiche come l'addizione e la moltiplicazione.

Come un messaggio trasformato da un codice per la correzione degli errori, la «risposta» di una dimostrazione probabilistica è distribuita su tutta la sua lunghezza, come lo sono gli eventuali errori. La dimostrazione viene controllata esaminandola in diversi punti e stabilendo se le risposte sono coerenti; all'aumentare del numero dei controlli, aumenta la certezza che l'argomentazione sia corretta. Laszló Babai dell'Università di Chicago, che ha messo a punto questo tipo di dimostrazioni due anni fa (insieme con Lance Fortnow, Carsten Lund e Mario Szegedy di Chicago e Leonid A. Levin dell'Università di Boston), le chiama «trasparenti». Manuel Blum di Berkeley, le cui ricerche hanno aperto la via al gruppo di Babai, propone invece il termine «olografiche».

Un avvenire incerto

Comunque le si chiami, queste dimostrazioni presentano alcuni inconvenienti pratici. Szegedy riconosce che trasformare una dimostrazione tradizionale in forma probabilistica è difficile e ne può risultare «qualcosa di molto più brutto e ingombrante». Per esempio una dimostrazione di 1000 righe può facilmente lievitare fino a 1000^3 (cioè un miliardo) righe. Tuttavia Szegedy sostiene che se lui e i suoi colleghi riuscissero a semplificare il procedimento di trasformazione, le dimostrazioni probabilistiche potrebbero diventare un metodo utile per verificare enunciati matematici e calcoli molto impegnativi (come quelli che portano alla dimostrazione del teorema dei quattro colori). «Il costo filosofico di questo metodo, così efficiente, è la perdita della certezza assoluta tipica di una dimostrazione euclidea» ha osservato Babai in un suo recente saggio, «ma se aveste dei dubbi, ve la sentireste di scommettere?».

Levin ritiene che sarebbe facile essere tratti in inganno da una tale scommessa, poiché anche un numero relativamente piccolo di controlli renderebbe trascurabile la probabilità di errore: uno diviso per il numero di particelle dell'universo. Levin rileva inoltre che perfino le più semplici dimostrazioni tradizionali sono soggette a dubbi di quest'ordine. «Per il principio d'indeterminazione di Heisenberg, nell'istante in cui scopro un errore il mio cervello potrebbe scomparire ed essere sostituito da un cervello che ritiene corretta la dimostrazione.»

Ronald L. Graham degli AT&T Bell Laboratories ipotizza che sia inevitabile una tendenza ad allontanarsi dalle brevi e chiare dimostrazioni tradizionali, che sono al di là di ogni ragionevole dubbio. «Può darsi» spiega «che le cose che è possibile dimostrare siano solo minuscole isole, eccezioni, in un vasto mare di risultati che non possono essere dimostrati dal solo pensiero umano». I matematici che cercano di solcare rotte non ancora segnate potrebbero essere costretti a dipendere sempre più da esperimenti, dimostrazioni probabilistiche e altre guide: «Forse non saremo più in grado di fornire dimostrazioni in senso classico».

Naturalmente, a mano a mano che gli studiosi diventeranno più dipendenti dal calcolatore, la matematica potrebbe fornire sempre meno soddisfazioni estetiche. «Sarebbe molto scoraggiante» osserva Graham «se a un certo punto io chiedessi a un calcolatore se l'ipotesi di Riemann è corretta e la macchina mi rispondesse: "Sì, ma non riuscirai mai a capire la dimostrazione"».

Al solo pensiero, senza dubbio, i tradizionalisti rabbrividiscono. Per il momento possono far quadrato intorno a eroi come Wiles, colui che ha sbaragliato l'ultimo teorema di Fermat, che rifugge da calcolatori, applicazioni e altri obbrobri. Ma se si vuol dare credito ai rapporti che giungono dal fronte dell'istruzione preuniversitaria, in futuro i Wiles potrebbero essere sempre più rari. Al Mathematical Sciences Research Institute di Berkeley, diretto da Thurston, si svolge una serie permanente di seminari con insegnanti delle scuole superiori per scoprire nuovi modi di attrarre gli studenti verso la matematica. Nel gennaio scorso la vicedirettrice dell'istituto, Lenore Blum, ha organizzato un seminario sul tema: «Sono superate le dimostrazioni di geometria nelle scuole superiori?».

I matematici hanno sostenuto che le dimostrazioni sono essenziali per garantire la verità di un risultato, ma gli insegnanti hanno obiettato che gli studenti considerano le tradizionali argomentazioni assiomatiche meno convincenti che non, per esempio, le dimostrazioni visive. Secondo gli atti del convegno, «la stragrande maggioranza degli insegnanti ha dichiarato che per lo più gli studenti di oggi (la generazione dei videogiochi e della televisione) non sono sensibili alle "dimostrazioni" e non ne capiscono l'importanza». Si notino le virgolette che accompagnano la parola «dimostrazioni».

BIBLIOGRAFIA

PETERSON IVARS, *Islands of Truth: A Mathematical Mystery Cruise*, W. H. Freeman and Company, 1990.

STEWART IAN, *The Problems of Mathematics*, Oxford University Press, 1992.

BARROW JOHN D., *Pi in the Sky: Counting, Thinking, and Being*, Oxford University Press, 1992.